The Saga of Reduced Factorizations of Elements of the Symmetric Group

by
A. M. Garsia

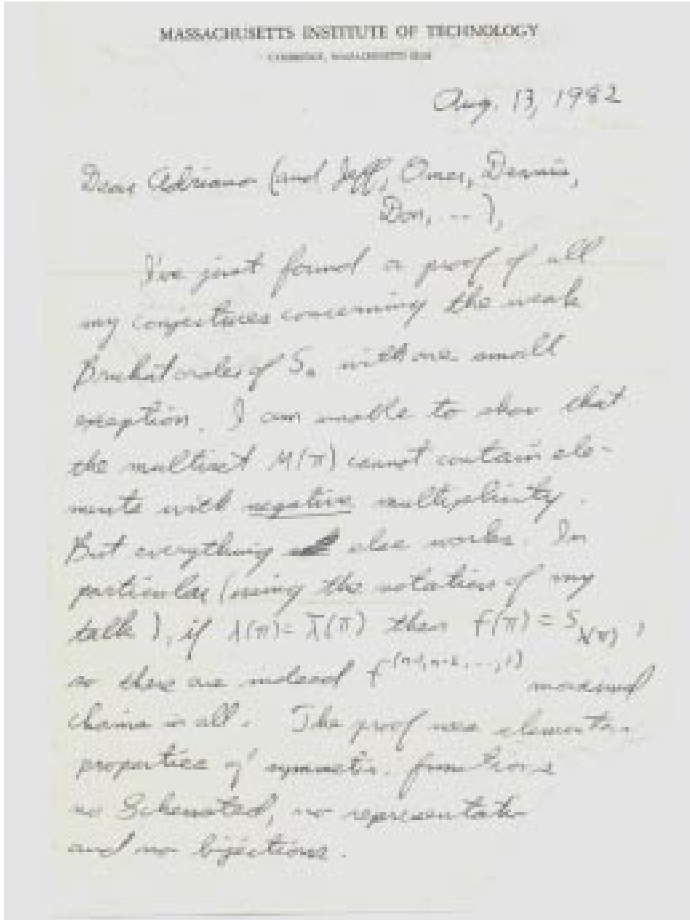
Abstract

These notes cover the contents of a series of lectures in a Topics in Algebraic Combinatorics course given at UCSD in Winter 2001. The initial effort was prompted by a desire to understand the connections between the theory of reduced decompositions started by the pioneering paper [] of R. Stanley and the theory of balanced tabloids studied by C. Green et al. [] [] However soon it appeared quite clear that a deeper understanding of the subject requires a parallel understanding of the Lascoux-Schützenberger theory of Schubert polynomials. These notes should offer a glimpse of the fascinating combinatorial connections between these theories. The presentation is generally self contained. The notes culminate with what should be a fairly lucid and illuminating proof of the Schur positivity of the Stanley symmetric functions.

TABLE OF CONTENTS

Introduction

In 1982 R. Stanley initiated the study of reduced decompositions of elements of S_n. Central to his work was a the introduction of a family of symmetric functions indexed by permutations. He conjectured these functions to be Schur positive and proved a number of their interesting properties including the enumeration of certain classes of reduced decompositions. Over the years that followed several works have appeared with different proofs of the Stanley conjecture by various methods which range from the purely combinatorial to the purely algebraic. Circa 1982 in a completely independent development Lascoux and Schützenberger founded the Theory of Schubert polynomials. Central to their study were some combinatorial consequences of a Pieri-like result for Schubert polynomials which they called “Monk's rule”. This led to the definition of a tree associated to every permutation $\sigma \in S_n$. Unbeknown to them at the time and to many even at the present time, the LS tree of a permutation is, in a sense that can be made precise, a purely combinatorial version of the Stanley symmetric function. Using this tree and several combinatorial properties of reduced decompositions, the Schur positivity of the Stanley symmetric function follows in a remarkably illuminating manner. In these notes we present the contents of a series of lectures in a Topics in Algebraic Combinatorics Course given at UCSD in Winter 2001. The material by no means covers all the aspects of the fascinating subject of reduced decompositions that have been developed over the last two decades. The choice of topics, limited by the time available, follows the taste of the author and what appeared to be a natural path through a luscious forest of remarkable combinatorial discoveries. We strived throughout to make our presentation as self-contained as possible. Some of the later proofs that appeared in the literature after the original papers are so elegant and simple that we were forced to reproduced them here almost verbatim. We claim no credit here for any of the results presented. This in only an expository work. Our main effort has been concentrated into providing a novel and illuminating way to develop the material. Our original stimulus for choosing this topic came from several exciting exchanges with Kevin Kadel a visitor at UCSD for the academic year 2000-2001. We also benefitted immensely from some of the insights he provided us in the study and developments connecting reduced decompositions to balanced tabloids.



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Aug. 13, 1982

Dear Adriano (and Jeff, Omer, Dennis, Don, ...),

I've just found a proof of all my conjectures concerning the weak Bruhat order of S_n with one small exception. I am unable to show that the multiset $M(\pi)$ cannot contain elements with negative multiplicity. But everything ~~else~~ else works. In particular (using the notation of my talk), if $\lambda(\pi) = \bar{\lambda}(\pi)$ then $f(\pi) = s_{\lambda(\pi)}$, so there are indeed $f^{(n-1, n-2, \ldots, 1)}$ maximal chains in all. The proof uses elementary properties of symmetric functions — no Schensted, no representation and no bijections.

(3)

The Saga of Reduced Factorizations of Elements of the Symmetric Group

by
A. M. Garsia

1. Reduced Factorizations

1.1 Notation

It is customary to interpret a permutation $\sigma \in S_n$ as a bijection of $\{1, 2, \ldots, n\}$ onto itself and we often write it in the form

$$\sigma \;=\; \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma_1 & \sigma_2 & \sigma_3 & \ldots & \sigma_n \end{pmatrix} \,,$$

meaning that σ_i is the image of i under σ. In this vein to compute the product $\theta \times \sigma$ we proceed from right to left and obtain

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \theta_{\sigma_1} & \theta_{\sigma_2} & \theta_{\sigma_3} & \ldots & \theta_{\sigma_n} \end{pmatrix} \;=\; \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \theta_1 & \theta_2 & \theta_3 & \ldots & \theta_n \end{pmatrix} \times \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma_1 & \sigma_2 & \sigma_3 & \ldots & \sigma_n \end{pmatrix}$$

Keeping this in mind, it will be convenient and economical with space to omit the first line and simply write

$$\sigma \;=\; \sigma_1\sigma_2\sigma_3\cdots\sigma_n$$

viewing σ as a word in the letters $1, 2, 3, \ldots, n$. Here and after we let s_i (for $1 \leq i \leq n-1$) represent the simple transposition

$$s_i = (i, i+1) = \begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & n \\ 1 & 2 & \cdots & i+1 & i & \cdots & n \end{pmatrix} \qquad 1.1.1$$

Note that multiplication of σ on the right by s_i results in the interchange of the elements σ_i, σ_{i+1}. Thus in our shorthand we may write

$$\sigma_1\sigma_2\cdots\sigma_{i+1}\sigma_i\cdots\sigma_n \;=\; \sigma_1\sigma_2\cdots\sigma_i\sigma_{i+1}\cdots\sigma_n \times s_i \;.$$

Let us recall that the number of inversions of σ is given by the sum

$$inv(sig) \;=\; \sum_{1\leq i<j\leq n} \chi(\sigma_i > \sigma_j)\;.$$

It is clear that right multiplication of σ by any simple transposition increases the number of inversions by one if $\sigma_i < \sigma_{i+1}$ and decreases it by one if $\sigma_i > \sigma_{i+1}$. Let us recall that an index i such that $\sigma_i > \sigma_{i+1}$ is called a "*descent*" of σ and correspondingly

$$D(\sigma) \;=\; \{1 \leq i \leq n-1 \,:\, \sigma_i > \sigma_{i+1}\}$$

is usually referred to as the "*descent set*" of σ. This given, if we want to express an element σ as a product of simple reflections the number of factors required should be at the very least $inv(\sigma)$. For this reason, $inv(\sigma)$ is often referred to as the "*length*" and briefly also denoted by $l(\sigma)$. Note that it is always possible (in fact in many ways) to express σ as a product of $l(\sigma)$ simple transpositions. To do this we simply start with $\sigma = \sigma^{(o)}$ and construct a sequence of permutations

$$\cdots \leftarrow \sigma^{(r+1)} \leftarrow \sigma^{(r)} \leftarrow \cdots \leftarrow \sigma^{(2)} \leftarrow \sigma^{(1)} \leftarrow \sigma^{(o)}$$

with $\sigma^{(r+1)} = \sigma^{(r)} \times s_i$ and where i is only chosen by the requirement that i be in the descent set of $\sigma^{(r)}$, that is $\sigma_i^{(r)} > \sigma_{i+1}^{(r)}$. Since this requirement assures that $l(\sigma^{(r+1)}) = l(\sigma^{(r)}) - 1$ the sequence will stop after exactly $l(\sigma)$ steps with $\sigma^{(l(\sigma))} = 123\cdots n$, (the identity permutation). In the display on the right we illustrate such a sequence for the permutation $\sigma = 35621784$. Here the labels on the right of the dividing line give the indices i for which the correspondind s_i was chosen. It should be apparent from this example that each time we have a variety of choices, (one for each element of the descent set of the current permutation).

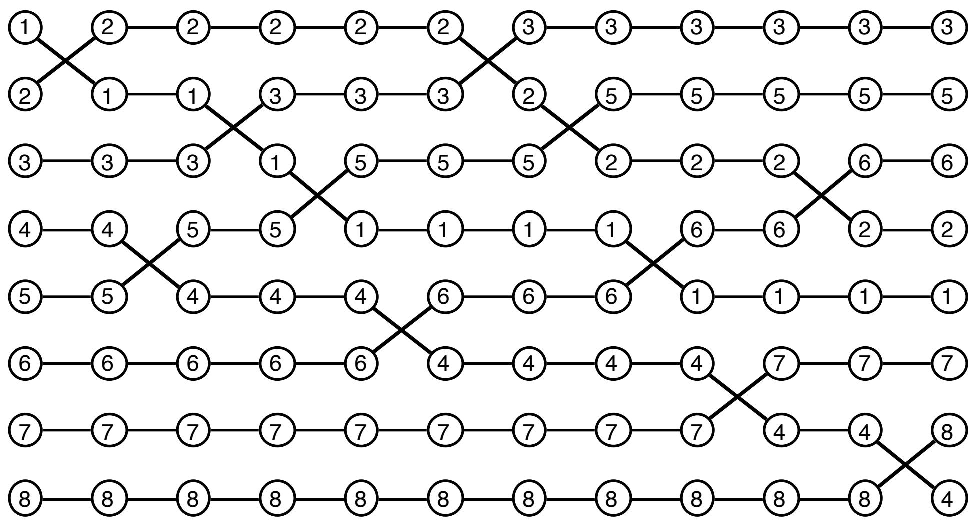
$$\begin{array}{c|c}
35621784 & \\
 & 7 \\
35621748 & \\
 & 3 \\
35261748 & \\
 & 6 \\
35261478 & \\
 & 4 \\
35216478 & \\
 & 2 \\
32516478 & \\
 & 1 \\
23516478 & \\
 & 5 \\
23514678 & \\
 & 3 \\
23154678 & \\
 & 2 \\
21354678 & \\
 & 4 \\
21345678 & \\
 & 1 \\
12345678 &
\end{array}$$

Factorizations of a permutation σ as a product of $l(\sigma)$ reflections are called "*reduced*" and the word in the letters $1, 2, \ldots, n-1$ giving the successive indices of the factors is called the "*reduced word* " corresponding to the factorization. Thus for the factorization above

$$35621784 = s_1 s_4 s_2 s_3 s_5 s_1 s_2 s_4 s_6 s_3 s_7 \qquad 1.1.2$$

the corresponding reduced word is 14235124637

Factorizations into simple reflections whether reduced or not are best studied by means of a line diagram which exhibits the trajectories of each of the labels $1, 2, \ldots, n$ as we proceed in our construction of the target permutation. In the display below we illustrate the diagram corresponding to the factorization illustrated above.



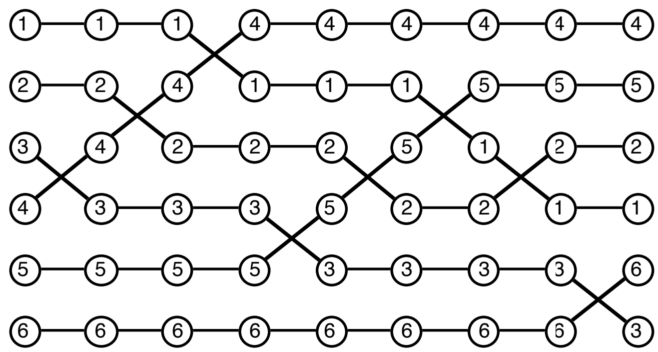
1.1.3

A close examination of this display reveals one fundamental property of diagrams corresponding to reduced factorizations:

for any pair of indices $1 \le i < j \le n$: the i-line and j-line cross at most once.

The reason for this is quite simple: once we interchange i and j, doing it again would decrease the number of inversions, and we never do that to get a reduced factorization.

We should mention that there is a systematic way of getting a reduced factorization for any permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n$. Starting from the identity permutation, we make first the interchanges that bring σ_1 to first position, then those that bring σ_2 to second position, then those that bring σ_3 to third position and so on until we reach σ. This is best understood by an example. In the next display we have illustrated this process applied to $\sigma = 452163$.



1.1.4

We thus obtain the factorization $452163 = s_3s_2s_1s_4s_3s_2s_3s_5$. It is easily seen that, in general, the resulting factorization will be of the form

$$\sigma \;=\; \prod_{i=1}^{n-1}\big(s_{a_i}s_{a_i-1}s_{a_i-2}\cdots s_{i+1}s_i\big) \tag{1.1.5}$$

with $a_i \geq i-1$ (note that $a_i = i-1$ must be included for the cases when the the corresponding factor should be taken equal to 1 (i.e. missing). Here and after these factorizations will be called "*canonical*". A moment's reflection should reveal that these observations yield the following basic identity

Theorem 1.1.1

$$\sum_{\sigma\in S_n}\sigma \;=\; \prod_{i=1}^{n-1}\big(1 + s_i + s_{i+1}s_i + s_{i+2}s_{i+1}s_i + \cdots + s_{n-1}\cdots s_{i+2}s_{i+1}s_i\,\big) \tag{1.1.6}$$

Proof

It should be understood that the factors in the right hand side of 1.1.6 are to be taken from left to right as i goes from 1 to $n-1$. This given, interpreting the left hand side as an element of the group algebra of S_n, then the identity simply asserts that each $\sigma\in S_n$ has a factorization of the form given in 1.1.5.

The following basic identities will play a fundamental role in the sequel, they are usually referred to as the "*Coxeter Relations*".

Proposition 1.1.1

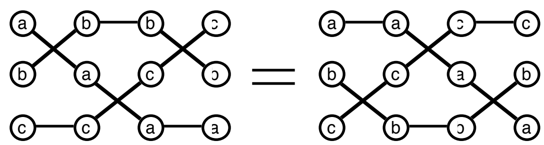
$$\begin{array}{lcl} 1) & s_i^2 = id & \forall \quad 1\leq i\leq n-1\,, \\ 2) & s_i\; s_{i+1}\, s_i = s_{i+1}\, s_i\; s_{i+1} & \forall \quad 1\leq i\leq n-1\,, \\ 3) & s_i\; s_j = s_j\, s_i & if \quad |i-j|\geq 2\,. \end{array} \tag{1.1.7}$$

Proof

The first and last follow immediately from the definitions of the s_i. The middle one just expresses the fact that the permutation

$$\theta_i = \begin{pmatrix} 1 & 2 & \cdots i & i+1 & i+2 & \cdots & n \\ 1 & 2 & \cdots i+2 & i+1 & i & \cdots & n \end{pmatrix}$$

has two reduced decompositions. We should also point out that the right hand side of 1.1.7 2) is in fact the canonical decomposition of θ_i. A visual understanding of this relation may also be provided by the following display



This is but an instance of the more general result which may be stated as follows

Theorem 1.1.2

We may pass from any reduced factorization to any other of a given permutation σ by a sequence of applications of identities 1.1.7 2) & 3). The inclusion of 1.1.7 1) is only necessary to pass from a non-reduced factorization of σ to a reduced one.

Proof

It is sufficient to show that we can pass from any factorization of σ to a canonical one. To this end our first step is to show that we may pass from any factorization which does not contain $s_1, s_2, \ldots, s_{i-1}$ to one which contains at most one occurrence of s_i. We can prove this by descent induction on i. Clearly the assertion is trivial for $i = n-1$. So let us assume that it is true for $i+1, i+2, \ldots, n-1$ and let W be a factorization which contains no occurrences of $s_1, s_2, \ldots, s_{i-1}$. Suppose W contains two occurrences of s_i and let us write it in the form

$$W = W_1 s_i W_2 s_i W_3 \tag{1.1.8}$$

with no occurrences of $s_1, s_2, \ldots, s_i$ in W_2. So by induction we change W_2 to a expression W_2' which contains no occurrences of s_{i+1} or one of the form

$$W_2' \;=\; W_{21}\, s_{i+1} W_{22}$$

with W_{21} and W_{22} not containing any occurrences of $s_1, s_2, \ldots, s_{i+1}$. In the first case, by successive uses of the Coxeter relations we can carry out the three transitions

$$W = W_1 s_i W_2 s_i W_3 \;\longrightarrow\; W_1 s_i W_2' s_i W_3 \;\longrightarrow\; W_1 s_i s_i W_2' W_3 \;\longrightarrow\; W_1 W_2' W_3 \;.$$

In fact, the second transition only needs successive uses of 1.1.7 3). Clearly, this case only occurs when W is not reduced.

In the other case, using the Coxeter relations we first carry out the transition

$$W = W_1 s_i W_2 s_i W_3 \;\longrightarrow\; W_1\, s_i\, W_{21}\, s_{i+1}\, W_{22}\, s_i\, W_3 \;.$$

Since W_{21} and W_{22} have only occurrences of s_j with $j > i+1$, by successive uses of 1.1.7 3) we can then carry out the transition

$$W_1\, s_i\, W_{21}\, s_{i+1}\, W_{22}\, s_i\, W_3 \quad\longrightarrow\quad W_1\, W_{21}\, s_i\, s_{i+1}\, s_i\, W_{22}\, W_3$$

and finally a use of 1.1.7 2) completes the sequence

$$\begin{aligned} W = W_1 s_i W_2 s_i W_3 \quad&\longrightarrow\; W_1\, s_i\, W_{21}\, s_{i+1}\, W_{22}\, s_i\, W_3 \quad\longrightarrow \\ &\longrightarrow\; W_1\, W_{21}\, s_i\, s_{i+1}\, s_i\, W_{22}\, W_3 \quad\longrightarrow\quad W_1\, W_{21}\, s_{i+1}\, s_i\, s_{i+1}\, W_{22}\, W_3 \end{aligned}$$

reducing by one the number of occurrence of s_i in W. Proceeding in this manner we can arrive at a point where either there is only one s_i left or none at all. This completes our induction. This given, starting from any factorization W, by means of the Coxeter relations we can eliminate altogether all the occurrences of s_1 or carry out the transition

$$W \quad\longrightarrow\quad W_1\, s_1\, W_2$$

with W_1 and W_2 containing no occurrences of s_1. By a further sequence of steps we can carry out one of the two transitions

$$W_1\, s_1\, W_2 \quad\longrightarrow\quad W_{11}\, s_2\, W_{12}\, s_1\, W_2 \qquad \textit{or} \qquad W_1\, s_1\, W_2 \quad\longrightarrow\quad W_1'\, s_1\, W_2$$

with no occurrences of s_1 or s_2 in W_{12} or W_1'. In each case successive uses of 1.1.7 3) will complete the succession of transitions

$$W \quad\longrightarrow\quad W_1\, s_1\, W_2 \quad\longrightarrow\quad W_{11}\, s_2\, W_{12}\, s_1\, W_2 \quad\longrightarrow\quad W_{11}\, s_2\, s_1\, W_{12} W_2$$

or

$$W \quad\longrightarrow\quad W_1\, s_1\, W_2 \quad\longrightarrow\quad W_1'\, s_1\, W_2 \quad\longrightarrow\quad s_1 W_1'\, W_2 \ .$$

Since there are no other occurrences of s_1 in either case and no ocurrences of s_1 or s_2 in W_{11} in the first case, we see that the pattern typical of a canonical factorization is beginning to emerge. Indeed the next step is to work on W_{11} and obtain one of the transitions $W_{11} \longrightarrow W_{111}\, s_3 W_{112}$ or $W_{11} \longrightarrow s_3 W_{112}$ with no occurrennces of s_1, s_2, s_3 in W_{112}. This gives the transitions

$$W \quad\longrightarrow\quad W_{11}\, s_2\, s_1\, W_{12} W_2 \quad\longrightarrow\quad W_{111}\, s_3 W_{112}\, s_2\, s_1\, W_{12} W_2 \quad\longrightarrow\quad W_{111} W_{112}\, s_3\, s_2\, s_1\, W_{12} W_2$$

or

$$W \quad\longrightarrow\quad W_{11}\, s_2\, s_1\, W_{12} W_2 \quad\longrightarrow\quad s_3 W_{112}\, s_2\, s_1\, W_{12} W_2 \quad\longrightarrow\quad s_3 s_2\, s_1\, W_{112}\, W_{12} W_2 \ .$$

We need not say any more here. The reader should have no difficulty understanding how this process can be continued to yield in the end a canonical decomposition of the permutation σ corresponding to the factorization W. To clear up any remaining uncertainties it may be appropriate to carry out the all the steps necessary in a particular instance. A good case in point is the factorization in 1.1.2. In the display below the labels on the right of the vertical line indicate which of the Coxeter relations are used in that particular transition the boxes appear as soon as one of the descent strings typical of canonical factorizations is formed.

$$\begin{array}{l|l}
1\ 4\ 2\ 3\ 5\ 1\ 2\ 4\ 6\ 3\ 7 & 3 \\
4\ 1\ 2\ 1\ 3\ 5\ 2\ 4\ 6\ 3\ 7 & 2 \\
4\ 2\ 1\ 2\ 3\ 5\ 2\ 4\ 6\ 3\ 7 & 3 \\
2\ 1\ 4\ 2\ 3\ 5\ 2\ 4\ 6\ 3\ 7 & 3 \\
2\ 1\ 4\ 2\ 3\ 2\ 5\ 4\ 6\ 3\ 7 & 2 \\
\boxed{2\ 1}\ \boxed{4\ 3\ 2}\ 3\ 5\ 4\ 6\ 3\ 7 & 3 \\
\boxed{2\ 1}\ \boxed{4\ 3\ 2}\ 5\ 3\ 4\ 3\ 6\ 7 & 2 \\
\boxed{2\ 1}\ \boxed{4\ 3\ 2}\ \boxed{5\ 4\ 3}\ \boxed{4}\ \boxed{6}\ \boxed{7} &
\end{array}$$

The main goal of these notes is to present some of the main results obtained in the description and enumeration of all reduced decompositions of any given permutation. Nevertheless, we should note at this point that, at least for small n, these reduced words can be constructed by computer in a relatively simple manner. This construction is based on the following identity.

Theorem 1.1.3

If for a given $\sigma \in S_n$, we denote by "$RED(\sigma)$" the collection of all words corresponding to reduced factorizations of σ then

$$\sum_{w\in RED(\sigma)} w \;=\; \sum_{i\in D(\sigma)} \sum_{w'\in RED(\sigma s_i)} w'i \qquad 1.1.9$$

Proof

It might be good to start by explaining the notation used in 1.1.9. To begin with the left hand side should be interpreted as the formal sum of all the elements of $R(\sigma)$. Thus to prove 1.1.9 we only have to show that each summand occurring in the left hand side occurs once and only once on the right hand side. Finally, we should note that the symbol "$w'i$" simply means the word obtained by appending the letter i to the word w'. Now note that if $W = W's_i$ is a reduced factorization of σ then we must necessarily have $\sigma_i > \sigma_{i+1}$ and W' will necessarily be a reduced factorization of $\sigma' = \sigma s_i$. This is because W' is a factorization of σ' and the number of its factors is $l(\sigma) - 1 = l(\sigma')$. Now if w is the word corresponding to W and w' is the word corresponding to W' we have $w = w'i$. This given we see that all $w \in RED(\sigma)$ do occur in the right hand side and they occur only once for the simple reason that each sum $\sum_{w'\in RED(\sigma s_i)} w'i$ consists of distinct words and different values of "i" yield different sums of words.

It will be instructive at this point to show how this identity can be translated into a MAPLE program. However, before implementing 1.1 we need a three auxiliary procedures "**sigact**", "**preds**", "**cocat**". The first has 2 input variables, an index i and a permutation σ . Then **sigact** returns the permutation $\sigma' = \sigma s_i$. The procedure **preds** takes a permutation σ as input and returns all the "*predecessors*" of σ, that is the collection

$$PRED(\sigma) \;=\; \{\sigma' \;:\; \sigma' = \sigma s_i \;\&\; \sigma_i > \sigma_{i+1}\} \qquad 1.1.10$$

Finally, **cocat** takes two input variables, an index s and a list of words L. Its output is the list of all words obtained by appending the index s to each word of L. These three procedures are given below

```
sigact:=proc(i,sig)
 local j,out;
 out:=[seq(sig[j],j=1..i-1),
       sig[i+1],sig[i],seq(sig[j],
            j=i+2..nops(sig))];
 end:
```

```
preds:=proc(sig)
 local n,out,i;
 n:=nops(sig);
 out:=NULL;
 for i from 1 to n-1 do
  if sig[i]>sig[i+1] then
  out:=out,[i,sigact(i,sig)];
   fi;
 od;
[out];
 end
```

```
cocat:=proc(s,L)
  local out,i,w;
  out:=NULL;
 for w in L do
cocat:=out,[op(w),s];
    od;
  out;
 end:
```

This given, the following procedure with input a permutation σ returns all the words corresponding to reduced factorizations of σ. It can be easily checked that it simply expresses in MAPLE almost verbatim the identity in 1.1.10.

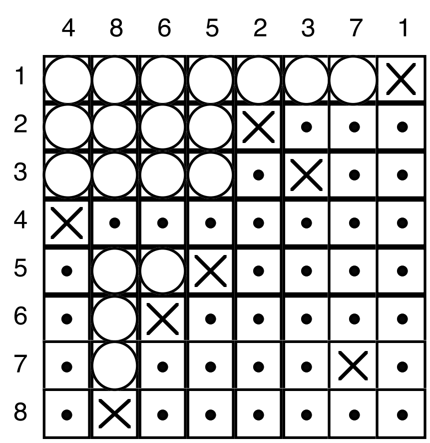
```
REDS:=proc(sig)
local prevs,out,i,s,m,tau,te,med;
 prevs:=preds(sig);
 if prevs=[] then
    out:=[[]];
   else
 te:=NULL;
  m:=nops(prevs);
   for i from 1 to m do
     s:=prevs[i][1];
     tau:=prevs[i][2];
    med:=cocat(s,REDS(tau));
 te:=te,med;
  od;
 out:=[te];
 fi;
 out;
 end;
```

Now a call of $\mathbf{REDS}([\mathbf{4},\mathbf{3},\mathbf{2},\mathbf{1}])$ yielded 16 reduced words as listed below.

123121	213231	312132	213213
121321	123212	132132	232123
212321	312312	321232	323123
231231	132312	231213	321323

1.1.11

We need to introduce a combinatorial structure which will play a crucial role in our further developments. Given a permutation $\sigma = \sigma_1\sigma_2\sigma_3\cdots\sigma_n$ we associate to it an $n \times n$ diagram with entries "$\bigcirc$", "X" or "$\bullet$", as follows. In column j and row σ_j we place an X. This done, in all the positions west or below this X we place an "$\bullet$". Finally when all the X's and the $\bullet$'s have been placed we fill the remaining positions with $\bigcirc$'s. The resulting figure will be referred to here and after as the "*Circle Diagram*" of the permutation σ. The display below gives the circle diagram of the permutation $\sigma = 48652371$.



1.1.12

Remark 1.2.1

We should note that each of the circles correspond to an inversion of σ. Indeed, from our construction of circle diagrams we will have a "$\bigcirc$" in position (i, j) if and only if the "X" in column j occurs below (i, j) and the "X" in row i occurs to the right of (i, j). This is equivalent to saying that $\sigma_j > i$ and $j' = \sigma_i^{-1} > j$, Thus this "$\bigcirc$" corresponds to the inversion $\sigma_j > \sigma_{j'}$.

1.2 The matrix approach

Note that the rearrangement

$$X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \quad \longrightarrow \quad X' = (x_4, x_8, x_6, x_5, x_2, x_3, x_7, x_1)$$

may simply be obtained by matrix multiplication. In fact, if we must have $X' = XM$ (interpreting X and X' as row vectors), then we are forced to take

$$M \;=\; \begin{pmatrix} 0&0&0&0&0&0&0&1\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&1&0\\ 0&1&0&0&0&0&0&0 \end{pmatrix} \qquad 1.2.1$$

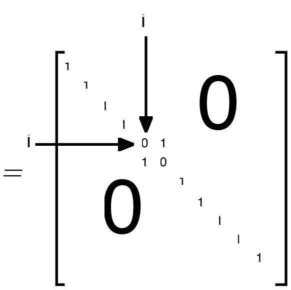
We clearly see that the positions of the ones in this matrix corresponds precisely to the positions of the $X's$ in the circle diagram of 48652371. More generally, the transition

$$X = (x_1, x_2, x_3, \ldots, x_n) \quad \longrightarrow \quad X' = (x_{\sigma_1}, x_{\sigma_2}, x_{\sigma_3}, \ldots, x_{\sigma_n})$$

can obtained be obtained by right multiplication of X by the matrix

$$M(\sigma) \;=\; \big\| \chi(i = \sigma_j) \big\|_{i,j=1}^{n}$$

We usually refer to $M(\sigma)$ as the "*permutation matrix*" corresponding to σ. Note then that the permutation matrix corresponding to the simple transposition $s_i = (i, i+1)$ of S_n may be schematically depicted as the $n \times n$ matrix

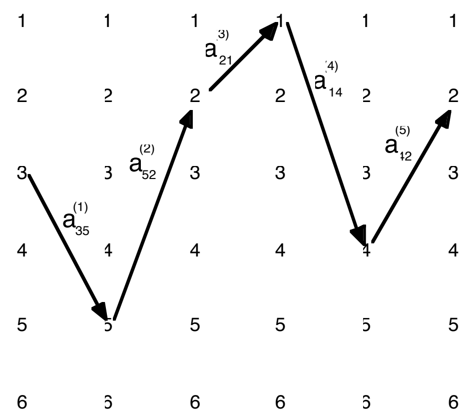
$M(s_i) =$ 

In other words, $M(s_i)$ has entries equal to one in positions $(i, i+1)$, $(i+1, i)$ and (j, j) for $j = 1, \ldots, i-1$ and $j = i+1, \ldots, n$, and all the remaining entries equal to zero.

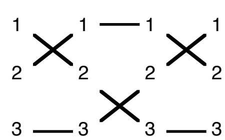
This enables us to view the line diagrams in 1.1.3 and 1.1.4 in a completely different light. Indeed, note that we may write the i,j-entry of the multiplication of $k{+}1$ matrices $A^{(r)} = \|a_{ij}^{(r)}\|_{i,j=1}^n$, $(r = 1,\ldots,k{+}1)$ in the form

$$(A^{(1)}A^{(2)}A^{(3)}\cdots A^{(k+1)})_{ij} \;=\; \sum_{i_1=1}^n \sum_{i_2=1}^n \sum_{i_3=1}^n \cdots \sum_{i_k=1}^n a_{i,i_1}^{(1)} a_{i_1,i_2}^{(2)} a_{i_2,i_3}^{(3)} \cdots a_{i_k,j}^{(k+1)} \;. \qquad 1.2.2$$

This expression has a very useful visualization. We depict a sequence of $k+2$ equally spaced columns, with nodes labelled $1,2,\ldots,n$ and view the sequence of indices $i{\rightarrow}i_1{\rightarrow}i_2{\rightarrow}\cdots{\rightarrow}i_k{\rightarrow}j$ as a path successively hitting the labels $i,i_1,i_2,\ldots,i_k,j$ as indicated below for the case $n=6$, $k=4$ and the sequence $3,5,2,1,4,2$. We also assign to the edge joining label i of column r to label j of column $r+1$ the "*weight*" $a_{i,j}^{(r)}$ and, correspondingly assign to any path a weight equal to the product of the weights of its edges. This given, we can then interpret the right hand side of 1.2.2 as the sum of the weights of all the paths joining label i of column 1 to label j of column $k+2$.



We shall here and after briefly refer to these displays as "*multiplication diagrams*". Clearly, the sum on the right hand side of 1.2.2 need only be carried out over the paths of weight $\neq 0$. This given, to further simplify these diagrams, we shall only depict edges $i{\rightarrow}j$ of weight $a_{ij}\neq 0$. In this manner the multiplication diagram of $M(s_1)M(s_2)M(s_1)$ reduces to



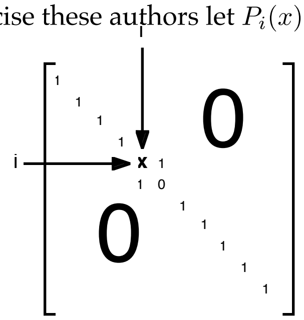
We can thus visualize the identity

$$M(s_1)M(s_2)M(s_1) \;=\; \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad 1.2.3$$

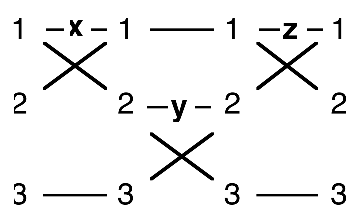
by computing each of the 9 i, j-entries in the product as a sum of weights of paths. Zero i, j-entries corresponding to the cases when there is no path joining i to j. Of course in this extremely simple case for any pair i, j either there is no path or there is only one of weight 1. This accounts for the right hand side of 1.2.3. Although we may not see it from this example, we will soon appreciate how powerful this imagery can be in understanding certain matrix identities. At any rate, we can now visualize the displays in 1.1.3 and 1.1.4 as instances of multiplication diagrams. In this manner we can use the display in 1.1.3 to obtain a visual understanding of the identity

$$M(s_1)M(s_4)M(s_2)M(s_3)M(s_5)M(s_1)M(s_2)M(s_4)M(s_6)M(s_3)M(s_7) = \begin{pmatrix} 0&0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&1\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0 \end{pmatrix}.$$

It develops that Kassel, Lascoux and Reutenauer [] discovered that by adding a single non-zero entry in each of the matrices $M(s_i)$ we can have the resulting product retain full information as to each of its factors and the order in which they occur. To be precise these authors let $P_i(x)$ (for a fixed n) be the $n \times n$ matrix

$P_i(x) =$ 

This given, it is easy to see that in the 3×3 case the product $P_1(x)P_2(y)P_1(z)$ may be represented by the multiplication diagram

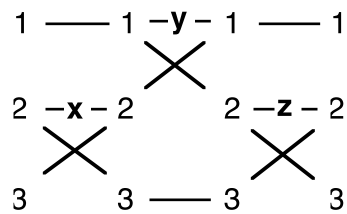


from which we derive that

$$P_1(x)P_2(y)P_1(z) = \begin{pmatrix} y+xz & x & 1\\ z & 1 & 0\\ 1 & 0 & 0 \end{pmatrix} \qquad 1.2.4$$

Here the $y + xz$ entry accounts for the fact that there are two paths joining 1 to 1. Namely, 1→1→1→1 and 1→2→2→1 of weights "xz" and "y" respectively.

Likewise from the diagram



we derive that

$$P_2(x)P_1(y)P_2(z) \;=\; \begin{pmatrix} y & z & 1 \\ x & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad 1.2.5$$

At this point it will be useful to note, for future reference, that combining 1.2.5 and 1.2.4 we obtain

$$P_1(x)P_2(y)P_1(z) \;=\; P_2(z)P_1(y+xz)P_2(x)$$

Similarly, in the $n \times n$ case, we derive that

$$P_i(x)P_{i+1}(y)P_i(z) \;=\; P_{i+1}(z)P_i(y+xz)P_{i+1}(x) \qquad \text{(for } i = 1,2,\ldots,n-1) \qquad 1.2.6$$

More generally, for a given reduced word $w = a_1a_2a_3\cdots a_l$ Kassel et al. do set in []

$$P_w(x_1,x_2,x_3,\ldots,x_l) \;=\; P_{a_1}(x_1)P_{a_2}(x_2)P_{a_3}(x_3)\cdots P_{a_l}(x_l)\ . \qquad 1.2.7$$

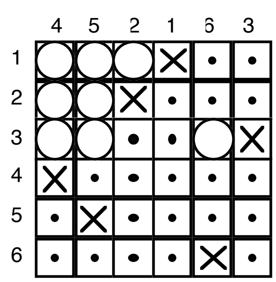
Our goal here is to fully understand the structure of this matrix. We shall begin by showing that in some cases its entries can be written down without any calculation. To be precise we have the following remarkable fact.

Theorem 1.2.1 (Kassell, et al.)

If w is the word of the canonical factorization of a permutation σ, then the matrix $P_w(x_1,x_2,x_3,\ldots,x_l)$ is simply obtained from the circle diagram of σ by replacing every "X" by a 1, every "•" by 0 and the "◯'s" by the variables $x_1,x_2,x_3,\ldots,x_l$ successively up the columns starting from the left most column and proceeding to the right.

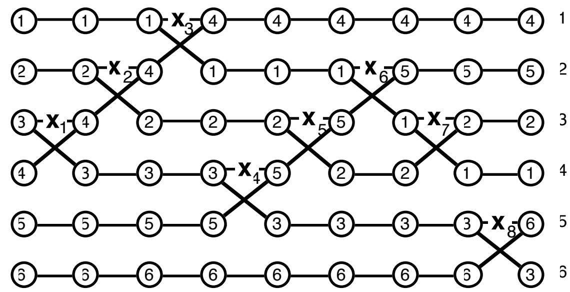
Proof

It will be good to start with a particular case. For instance, for the canonical factorization of $\sigma = 452163$, illustrated in 1.1.4, this construction yields



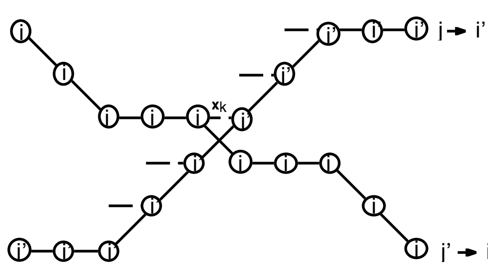
$$\longrightarrow \quad P_{32143235}(x_1,x_2,\ldots,x_8) = \begin{pmatrix} x_3 & x_6 & x_7 & 1 & 0 & 0 \\ x_2 & x_5 & 1 & 0 & 0 & 0 \\ x_1 & x_4 & 0 & 0 & x_8 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \qquad 1.2.8$$

To visualize the mechanism that produces this result we resort to the multiplication diagram corresponding to the product that yields $P_{32143235}(x_1,x_2,\ldots,x_8)$. Now it is not difficult to see that this diagram can be simply obtained by adding edges with weights $x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8$ to the display in 1.1.4, as indicated below



To calculate the $3,2$-entry in $P_{32143235}(x_1,x_2,\ldots,x_n)$ using this diagram we locate all the paths that join 3 to 2. We see that there is only one such path. This is obtained by following the 3-line until it meets the edge labled x_4 then traverse this edge and then follow the 5-line untill the end. This gives that the $3,2$-entry is x_4. Now we should clearly see why the entries in positions $(3,1)$, $(2,1)$ and $(1,1)$ turn out to be x_1,x_2,x_3 respectively. This is simply because as we bring 4 to first position by the transpositions s_3,s_2,s_1, in the product diagram corresponding to $P_3(x_1)P_2(x_2)P_1(x_3)$ the horizontal edges with weights x_1,x_2,x_3 open up three paths respectively joining 3 to 1, 2 to 1 and 1 to 1. Similarly in the portion of the diagram corresponding to the factors $P_4(x_4)P_3(x_5)P_2(x_6)$ the horizontal edges with weights x_4,x_5,x_6 open up three paths respectively joining 3 to 2, 2 to 2 and 1 to 2. That accounts for x_4,x_5,x_6 landing in positions $(3,2)$, $(2,2)$, $(1,2)$, of the resulting matrix. Similar reasoning accounts for the positions of x_7 and x_8. To establish the result in the general case, we have three crucial observations:

First, we note that because in a canonical factorization, we bring the elements σ_1,$\sigma_2,\sigma_3\ldots$ to their positions successively one at the time, as we bring σ_j to the j^{th} in steps $k,k+1,k+2,\ldots,k+r$ the edges with weights $x_k,x_{k+1},x_{k+2},\ldots,x_{k+r}$ are all above the σ_j-line. This given, when a path in the multiplication diagram traverses one of these edges it will then be forced to follow the σ_j-line to its end and therefore it will never be able to traverse any other x-weighted edge. This shows that for any pair (i,j) there is no path joining i to j, or a single path. In the latter case the path starts with the i-line and either it never traverses one of the x-weighted edges thereby following the i-line all the way to the end (here $i=\sigma_j$ and the i,j-entry is "1") or it traverses an x-weighted edge and then it must continue along the $i'=\sigma_j^{-1}$-line all the way to the end (see figure below)



If the crossing occurs at step k then the weight of the edge is x_k and the i,j-entry will be x_k.

Second, we note that in the latter case, $\sigma_{j'} = i$ (see figure above) with $j' > j$ and $\sigma_j = i' > i$ imply that the i, j-position is precisely a "◯"-position in the circle diagram of σ.

Finally, if the weights of horizontal x-labelled edges that touch the σ_j-line are successively $x_i, x_{i+1}, \ldots, x_{i+r}$ then these weights will necessarily land in the "◯"-positions of the j^{th} column of the resulting matrix. This completes our proof.

Remark 1.2.2

We have shown above that if the k^{th} transposition in our reduced expression interchanges i with $i' = \sigma_j$ then the variable x_k will appear in the i, j-entry of the resulting matrix. If we review the argument we can easily see that this particular conclusion did not use the fact that there we were dealing with a canonical factorization. However, in the general case, as we shall see, there will also be other paths joining i to j and they will contribute further terms to the i, j-entry of the resulting matrix. Keeping in mind this fact we can prove the following remarkable property of the matrices $P_w(x_1, x_2, \ldots, x_l)$.

Theorem 1.2.2 (Kassel, et al.)

Let σ be a permutation of length l and let $\mathcal{J} = (x_i x_j : 1 \leq i < j \leq l)$ be the ideal in the polynomial ring $\mathbf{Q}[x_1, x_2, \ldots, x_l]$ generated by the products $x_i x_j$. Then for any $w \in RED(\sigma)$ the matrix $P_w(x_1, x_2, \ldots, x_l)$ modulo J may be obtained from the circle diagram of σ by replacing every "X" by a 1, every "•" by 0 and the "◯'s" by a permutation of the variables $x_1, x_2, x_3, \ldots, x_l$. More precisely, if $w = a_1, a_2, \ldots, a_l$ then the "◯" in position (i, j) is to be replaced by x_k if the transposition s_{a_k} interchanges i with σ_j.

Proof

Recall that we can pass from w to the canonical factorization w_o of σ by a succession of applications of the relations 2) and 3) of 1.1.7. Now from 1.2.6 we deduce that

$$P_i(x)P_{i+1}(y)P_i(z) \cong P_i(z)P_{i+1}(y)P_i(x) \pmod{\mathcal{J}} \quad \text{for } i = 1, \ldots, l \qquad 1.2.9$$

and we clearly have

$$P_i(x)P_j(y) = P_j(y)P_i(x) \quad \text{for } |j - i| \geq 2 \quad . \qquad 1.2.10$$

Thus if we use the same relations that bring us from w to w_o to the product

$$P_w(x_1, x_2, \ldots, x_l) = P_{a_1}(x_1)P_{a_2}(x_2)P_{a_3}(x_3) \cdots P_{a_l}(x_l) \ ,$$

we see that the relations in 1.2.9 and 1.2.10 will yield us an identity of the form

$$P_w(x_1, x_2, \ldots, x_l) \cong P_{w_o}(x_{\theta_1}, x_{\theta_2}, \ldots, x_{\theta_l}) \pmod{\mathcal{J}}$$

with $\theta_1, \theta_2, \ldots, \theta_l$ a permutation of $1, 2, \ldots, n$. This given, our assertions follow from Theorem 1.2.1 and Remark 1.2.1.

It will be worthwhile to illustrate this argument by working on a specific example. For this we take $\sigma = 615243$ and the word $w = 453243251 \in RED(\sigma)$. In the display below we give the sequence of steps that transform 453243251 into the canonical factorization 543215435 of σ. On the right of the vertical line we have indicated the transformation we carried out from one step to the next.

$$
\begin{array}{ccccccccc|l}
P_4(x_1) & P_5(x_2) & P_3(x_3) & P_2(x_4) & P_4(x_5) & P_3(x_6) & P_2(x_7) & P_5(x_8) & P_1(x_9) & 51 \to 15 \\
P_4(x_1) & P_5(x_2) & P_3(x_3) & P_2(x_4) & P_4(x_5) & P_3(x_6) & P_2(x_7) & P_1(x_9) & P_5(x_8) & 2432 \to 4232 \\
P_4(x_1) & P_5(x_2) & P_3(x_3) & P_4(x_5) & P_2(x_4) & P_3(x_6) & P_2(x_7) & P_1(x_9) & P_5(x_8) & 232 \to 323 \\
P_4(x_1) & P_5(x_2) & P_3(x_3) & P_4(x_5) & P_3(x_7) & P_2(x_6) & P_3(x_4) & P_1(x_9) & P_5(x_8) & 31 \to 13 \\
P_4(x_1) & P_5(x_2) & P_3(x_3) & P_4(x_5) & P_3(x_7) & P_2(x_6) & P_1(x_9) & P_3(x_4) & P_5(x_8) & 343 \to 434 \\
P_4(x_1) & P_5(x_2) & P_4(x_7) & P_3(x_5) & P_4(x_3) & P_2(x_6) & P_1(x_9) & P_3(x_4) & P_5(x_8) & 3421 \to 3214 \\
P_4(x_1) & P_5(x_2) & P_4(x_7) & P_3(x_5) & P_2(x_6) & P_1(x_9) & P_4(x_3) & P_3(x_4) & P_5(x_8) & 454 \to 545 \\
P_5(x_7) & P_4(x_2) & P_5(x_1) & P_3(x_5) & P_2(x_6) & P_1(x_9) & P_4(x_3) & P_3(x_4) & P_5(x_8) & 5321 \to 3215 \\
P_5(x_7) & P_4(x_2) & P_3(x_5) & P_2(x_6) & P_1(x_9) & P_5(x_1) & P_4(x_3) & P_3(x_4) & P_5(x_8) &
\end{array}
\qquad 1.2.11
$$

This shows that modulo the ideal $\mathcal{J} = \big(x_i x_j : 1 \leq i < j \leq 9\big)$ we have

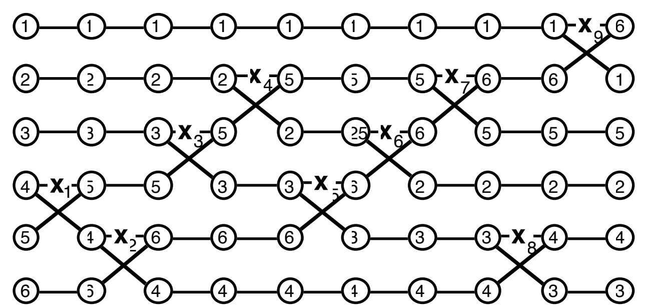
$$P_{453243251}(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) \cong P_{543215435}(x_7, x_2, x_5, x_6, x_9, x_1, x_3, x_4, x_8,) \qquad 1.2.12$$

Since $s_5 s_4 s_3 s_2 s_1 s_5 s_4 s_3 s_5$ is the canonical factorization of $\sigma = 615243$ we can follow the recipe given by Theorem 1.2.1 and obtain

$$P_{453243251}(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) \cong \begin{pmatrix} x_9 & 1 & 0 & 0 & 0 & 0 \\ x_6 & 0 & x_4 & 0 & 0 & 0 \\ x_5 & 0 & x_3 & 0 & x_8 & 1 \\ x_2 & 0 & x_1 & 0 & 1 & 0 \\ x_7 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad 1.2.13$$

Remark 1.2.3

We should note that the effect of working in the quotient ring $\mathbf{Q}[x_1, x_2, \ldots, x_l]/\mathcal{J}$ is to kill all contributions to the matrix $P_w(x_1, x_2, \ldots, x_l)$ coming from paths that traverse more than one of the x-weighted edges. In fact we can easily see from the adjoining product diagram that the if we do not kill all monomials of degree 2 the resulting matrix is



$$P_{453243251}(x_1, x_2, \ldots, x_9) \;\; = \;\; \begin{pmatrix} x_9 & 1 & 0 & 0 & 0 & 0 \\ x_6 + x_4 x_7 & 0 & x_4 & 0 & 0 & 0 \\ x_5 + x_3 x_7 & 0 & x_3 & 0 & x_8 & 1 \\ x_2 + x_1 x_7 & 0 & x_1 & 0 & 1 & 0 \\ x_7 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Our next goal is to show that we can produce equivalences such as in 1.2.12 by working directly with the final matrices, rather than by acting on the factors. To state and prove this result we need to make some definitions and establish some auxiliary propositions. To begin let us denote by $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ the matrix we obtain when we compute the entries of $P_w(x_1, x_2, \ldots, x_l)$ mod $\mathcal{J}$. We shall also refer to $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ as the "*linear part*" of $P_w(x_1, x_2, \ldots, x_l)$. For given indices $j_1 < j_2 < \cdots < j_k$, let us denote by $P_w^{\mathcal{J}}[j_1, j_2, \ldots, j_k]$ the $k \times k$ submatrix of $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ contained in columns $j_1, j_2, \ldots, j_k$ and rows $\sigma_{j_1}, \sigma_{j_2}, \ldots, \sigma_{j_k}$. Note that if $k = 3$ and $\sigma_{j_1} > \sigma_{j_2} > \sigma_{j_3}$ then the submatrix $P_w^{\mathcal{J}}[j_1, j_2, j_3]$ will be of the form

$$P_w^{\mathcal{J}}[j_1, j_2, j_3] \;=\; \begin{pmatrix} y & z & 1 \\ x & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

This given, we shall call a "*3-Coxeter transition for k*" in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ a replacement of the form

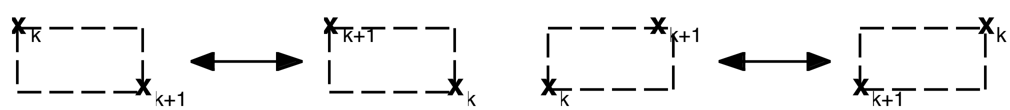
$$\begin{pmatrix} x_{k+1} & x_{k+2} & 1 \\ x_k & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \longleftrightarrow \quad \begin{pmatrix} x_{k+1} & x_k & 1 \\ x_{k+2} & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad 1.2.14$$

More precisely, such a transition consists in locating three indices $j_1 < j_2 < j_3$ such that the submatrix $P_w^{\mathcal{J}}[j_1, j_2, j_3]$ is of one of the forms given in 1.2.14. This done, the 3-Coxeter transition consists in replacing one form by the other form in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$.

In the same vein, a "*2-Coxeter transition on k*" is the exchange of x_k and x_{k+1} when

$$\underline{x_k \;\; \textit{and} \;\; x_{k+1} \;\; \textit{are not in the same row or column.}}$$

Thus this Coxeter transition carries out one of the following 4 possible exchanges in the matrix $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$:



Proposition 1.2.1

Let $w = a_1, a_2, \cdots a_l$ be a reduced word and let

$$a_k = i \quad , \quad a_{k+1} = i + 1 \quad , \quad a_{k+2} = i \; .$$

Let $w' = a'_1, a'_2, \cdots a'_l$ be the same as w except in positions $k, k+1, k+2$ where we have

$$a'_k = i + 1 \quad , \quad a'_{k+1} = i \quad , \quad a'_{k+2} = i + 1$$

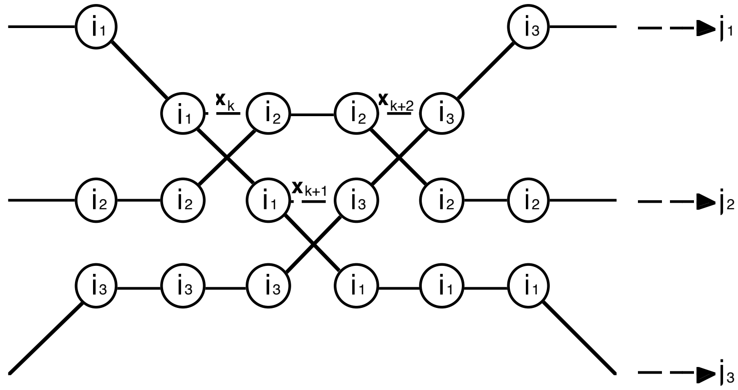
Then the matrix $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ is simply obtained from $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ by making a 3-Coxeter transition on k.

Proof

We have

$$P_w(x_1, x_2, \ldots, x_l) \;=\; P_{a_1}(x_1) \cdots P_i(x_k) P_{i+1}(x_{k+1}) P_i(x_{k+2}) \cdots P_{a_l}(x_l) \; .$$

Under this assumption, the portion of the diagram that contains the edges of weights x_k, x_{k+1} and x_{k+2} will necessarily be of the form given below with the x_k, x_{k+1} and x_{k+2} edges at heights i, $i+1$ and i respectively.

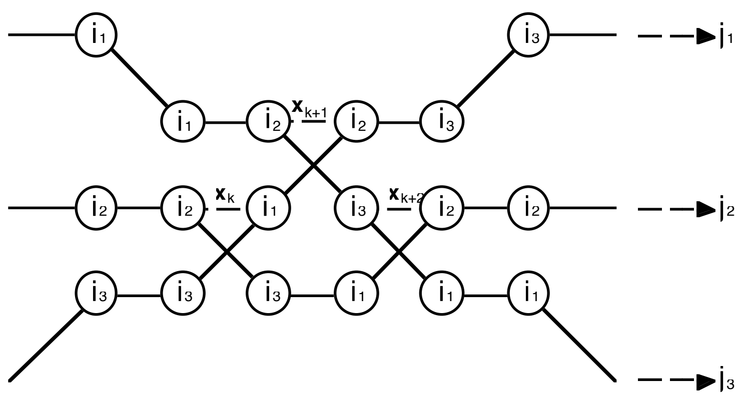


Indeed, if it is i_1-line and the i_2-line that cross at thew k^{th} step, and if it is the i_3-line that the i_1-line crosses at the $k+1^{st}$ step then the i_2 and i_3 lines will necessarily cross at the $k+2^{nd}$ step. Since, in the line diagram of a reduced decomposition, any two labelled lines cross only once, we will have $i_1 < i_2 < i_3$ and the i_3, i_2 and i_1 lines must respectively end up at levels $j_1 < j_2 < j_3$ as indicated in the figure. Of course this means that $\sigma_{j_1} = i_3$, $\sigma_{j_2} = i_2$ and $\sigma_{j_3} = i_1$

Using this diagram and the recipe given by Theorem 1.2.3, we can easily derive that the submatrix $P_w^{\mathcal{J}}[j_1, j_2, j_3]$ must be precisely as given below

$$P_w^{\mathcal{J}}[j_1, j_2, j_3] \;=\; \begin{pmatrix} x_{k+1} & x_k & 1 \\ x_{k+2} & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad . \tag{1.2.16}$$

Note next that if the portion of the product diagram of $P_{a_1}(x_1)P_{a_2}(x_2)\cdots P_{a_l}(x_l)$ given above, is replaced by the portion given below



what we get is precisely the multiplication diagram we can use to compute the matrix

$$P_{w'}(x_1, x_2, \ldots, x_l) \;=\; P_{a'_1}(x_1) \cdots P_{i+1}(x_k) P_i(x_{k+1}) P_{i+1}(x_{k+2}) \cdots P_{a'_l}(x_l) \;.$$

On the other hand, the relation in 1.2.7 (modulo $\mathcal{J}$) gives

$$P_{i+1}(x_k) P_i(x_{k+1}) P_{i+1}(x_{k+2}) \;\cong\; P_i(x_{k+2}) P_{i+1}(x_{k+1}) P_i(x_k) \qquad (\text{mod } \mathcal{J})$$

This means that we also have

$$P^{\mathcal{J}}_{w'}(x_1, x_2, \ldots, x_l) \;\cong\; P_{a_1}(x_1) \cdots P_i(x_{k+2}) P_{i+1}(x_{k+1}) P_i(x_k) \cdots P_{a_l}(x_l) \;\cong\; P^{\mathcal{J}}_{w}(x_1, \cdots, x_{k+2}, x_{k+1}, x_k, \ldots, x_l)$$

In other words $P^{\mathcal{J}}_{w'}(x_1, x_2, \ldots, x_l)$ is obtained from $P^{\mathcal{J}}_{w}(x_1, x_2, \ldots, x_l)$ by interchanging x_k with x_{k+2}. However, in view of 1.2.16 this is precisely a 3-Coxeter transition on k.

It is important to know at this point how the matrix $P^{\mathcal{J}}_{w}(x_1, x_2, \ldots, x_l)$ changes as we increase or decrease the number of factors. It develops that these changes can be carried out by a very simple recipe. More precisely we have

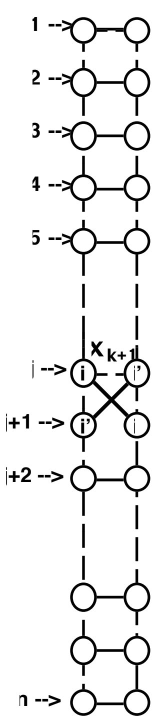
Proposition 1.2.2

Let $w = a_1 a_2 \cdots a_k \in RED(\sigma)$, and let $\sigma_j < \sigma_{j+1}$ so that $w' = a_1 a_2 \cdots a_k j \in RED(\sigma \times s_j)$, then the transition

$$P^{\mathcal{J}}_{w}(x_1, x_2, \ldots, x_k) \quad \longrightarrow \quad P^{\mathcal{J}}_{w'}(x_1, x_2, \ldots, x_{k+1})$$

is simply obtained by interchanging columns j and $j+1$ of $P^{\mathcal{J}}_{w}(x_1, x_2, \ldots, x_k)$ and then changing the (σ_j, j)-entry of the resulting matrix to "x_{k+1}".

Proof



For convenience let $\mathcal{M}_w$ and $\mathcal{M}_{w'}$ denote the multiplications diagrams corresponding to w and w' and let $\mathcal{M}_{w'/w}$ denote the the last two columns we have to add to $\mathcal{M}_w$ to get $\mathcal{M}_{w'}$. Since by our assumptions we have

$$P_{w'}(x_1, x_2, \ldots, x_{k+1}) \;=\; P_w(x_1, x_2, \ldots, x_k) \times P_j(x_{k+1}) \;,$$

the diagram $\mathcal{M}_{w'/w}$ will necessarily be as depicted in the the adjacent figure. We have also set there $i = \sigma_j$ and $i' = \sigma_{j+1}$. Now note that, when $s \neq j$ or $s \neq j+1$, to compute an r, s entry in the matrix $P_{w'}(x_1, x_2, \ldots, x_{k+1})$ we simply follow the same paths as for the computation of the r, s entry of $P_w(x_1, x_2, \ldots, x_k)$ up to the first column of $\mathcal{M}_{w'/w}$ and then proceed to the second column of $\mathcal{M}_{w'/w}$ traversing the horizontal edge at level s. This yields that the s^{th} columns of $P_{w'}(x_1, x_2, \ldots, x_k)$ and $P_w(x_1, x_2, \ldots, x_k)$ are identical. Similarly we see that, to compute an $r, j+1$-entry of $P_{w'}(x_1, x_2, \ldots, x_k)$, we must follow a path of $\mathcal{M}_w$ that goes from r to level j and then drop down to level $j+1$ by following the last step of the i-line in $\mathcal{M}_{w'/w}$. This causes the $j+1^{st}$ column of $P^{\mathcal{J}}_{w'}(x_1, x_2, \ldots, x_{k+1})$ to be identical with the j^{th} column of $P^{\mathcal{J}}_{w}(x_1, x_2, \ldots, x_k)$. To compute the r, j-entry of $P^{\mathcal{J}}_{w'}(x_1, x_2, \ldots, x_{k+1})$ we have two sets of paths. Those which in $\mathcal{M}_w$ go from r to level j and continue in $\mathcal{M}_{w'/w}$ horizontally by traversing the x_{k+1}-weighted edge (see figure), and those which in $\mathcal{M}_w$ go from r to level $j+1$ and then climb up to level j by following the last step of the i'-line in $\mathcal{M}_{w'/w}$.

However from the first set of paths only the i-line survives in computation mod $\mathcal{J}$. The reason for this is,except for the i-line, all the other paths have contributed an x-entry in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_k)$ and the continuation across the the x_{k+1}-weighted edge will make their weight a product of $x's$ and therefore equal to zero mod $\mathcal{J}$. On the other hand the i-line in $\mathcal{M}_w$ followed by the x_{k+1}-weighted edge will contribute an x_{k+1} to the i, j-entry of $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_{k+1})$. Now a path in $\mathcal{M}_w$ from second set that goes from an $r \neq i$ to level $j+1$, yields the $r, j+1$-entry in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_k)$ and will cause this entry to move to the r, j position in $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_{k+1})$ as it climbs to level j in $\mathcal{M}_{w'/w}$.

We have now accounted for all but the i, j-entry in $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_{k+1})$. The possibility remains that x_{k+1} may not be the only term there because of some path from second set that went from i to level $j+1$ in $\mathcal{M}_w$. However note that since $\sigma_j = i$ there is no "$\bigcirc$" or "X" in position $i, j+1$ in the circle diagram of σ so the i, j-entry in $P_w(x_1, x_2, \ldots, x_k)$ is necessarily zero and therefore there is no path in $\mathcal{M}_w$ that joins i to $j+1$. Thus the i, j-entry of $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_{k+1})$ must be x_{k+1} precisely as asserted.

In the display below we illustrate the sequence of transitions corresponding to the reduced word $w = 24534231 \in RED(516324)$.

$$\begin{pmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{pmatrix} -2\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_1&1&0&0&0\\ 0&1&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{pmatrix} -4\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_1&1&0&0&0\\ 0&1&0&0&0&0\\ 0&0&0&x_2&1&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&1 \end{pmatrix} -5\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_1&1&0&0&0\\ 0&1&0&0&0&0\\ 0&0&0&x_2&x_3&1\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0 \end{pmatrix}$$

$$-3\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_1&x_4&1&0&0\\ 0&1&0&0&0&0\\ 0&0&x_2&0&x_3&1\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0 \end{pmatrix} -4\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_1&x_4&x_5&1&0\\ 0&1&0&0&0&0\\ 0&0&x_2&x_3&0&1\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0 \end{pmatrix} -2\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_4&x_1&x_5&1&0\\ 0&x_6&1&0&0&0\\ 0&x_2&0&x_3&0&1\\ 0&1&0&0&0&0\\ 0&0&0&1&0&0 \end{pmatrix} \qquad 1.2.17$$

$$-3\rightarrow \begin{pmatrix} 1&0&0&0&0&0\\ 0&x_4&x_5&x_1&1&0\\ 0&x_6&x_7&1&0&0\\ 0&x_2&x_3&0&0&1\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0 \end{pmatrix} -1\rightarrow \begin{pmatrix} x_8&1&0&0&0&0\\ x_4&0&x_5&x_1&1&0\\ x_6&0&x_7&1&0&0\\ x_2&0&x_3&0&0&1\\ 1&0&0&0&0&0\\ 0&0&1&0&0&0 \end{pmatrix}$$

Since each of these matrix transitions can be reversed, an immediate corollary of Proposition 1.2.3 is that the word w can be reconstructed from $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$. To do this we simply carry out the illustrated process in reverse. In particular we obtain thus a proof that the matrix $P_w(x_1, x_2, \ldots, x_l)$ is completely determined by its linear part. Now it develops that there is an even simpler way, in fact a recipe, for recovering w from $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$. This result can be stated as follows.

Theorem 1.2.3 (C. Greene et Al [])

Let $w = a_1 a_2 \cdots a_l$, and for each $k \in [1, l]$ let c_k denote the number of x_s with $s > k$ that are directly NORTH or SOUTH of x_k in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ and let r_k be the number of x_s with $s > k$ that are directly WEST. This given, if x_k is in column j_k of $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ we necessarily have

$$a_k \quad = \quad j_k + c_k - r_k \tag{1.2.18}$$

Proof

It follows from Proposition 1.2.2 and it is easy to see from the process in 1.2.17 that if $a_k = j$ then x_k lands in column j at the moment it is inserted. However, as the process of construction of $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ continues, its column changes. Nevertheless we can easily keep track of what happens. To begin we see that every time an x_s with $s > k$ gets inserted in the column of x_k the column number of x_k decreases by one. On the other hand note that if an x_s with $s > k$ gets inserted in the row of x_s this will necessarily take place EAST of x_k, because to the right of x_s we place a 1 and there is nothing but zeros in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ to the right of any 1's. Now the only time when such an x_s passes to the WEST of x_k is when x_s is immediately to the right of x_k and their columns are interchanged. This causes the column number of x_k to increase by one at that time. Putting all this together we derive that when the transition process terminates we will find x_k in column j_k with

$$j_k = a_k + r_k - c_k \ .$$

This proves 1.2.18.

Proposition 1.2.3

Let $w = a_1, a_2, \cdots a_l$ be a reduced word and let

$$a_k = r \quad \text{and} \quad a_{k+1} = s \quad \text{with } |r - s| \geq 2 \ . \tag{1.2.19}$$

Let $w' = a'_1, a'_2, \cdots a'_l$ be the same as w except in positions $k, k+1$ where we have

$$a'_k = s \quad \text{and} \quad a'_{k+1} = r \tag{1.2.20}$$

Then the matrix $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ is simply obtained from $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ by making a 2-Coxeter transition on k.

Proof

Let $M_w^{(h)}$ for a moment denote the matrix obtained after h steps in the construction process that yields $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$. Likewise let $M_{w'}^{(h}$ be the matrix obtained after h steps in the construction process that yields $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$. This given, from Proposition 1.2.2 and 1.2.19 it follows that x_k and x_{k+1} will respectively be in columns r and s of $M_w^{(k+1)}$. It is also clear that x_{k+1} is not inserted in the same row as x_k because immediately to the right of x_k in $M_w^{(k)}$ there is a 1.

Now note that since the two columns involved in the insertion of x_k do not overlap with the two columns involved in the insertion of x_{k+1} we can easily see that $M_{w'}^{(k+1)}$ will necessarily be identical with $M_w^{(k+1)}$ except that the positions of x_k and x_{k+1} are interchanged. Consequently, during the remaining part of the insertion processes yielding $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ and $P_{w'}^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ we shall have that $M_w^{(h)}$ and $M_{w'}^{(h)}$ will remain related by a 2-Coxeter transition and it will be so the end as well proving our assertion.

We are now finally in a position to establish the following basic result.

Theorem 1.2.3

For any two reduced words w_1 and w_2 of a permutation σ of length l we can find a sequence of Coxeter transitions that transform $P_{w_1}^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ into $P_{w_2}^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$.

Proof

By Theorem 1.1.2 we can pass from w_1 to w_2 by a sequence of applications of identities 1.1.17 2) and 3). But now from Propositions 1.1.2 and 1.1.3 we derive that 1.1.17 2) will cause a 3-Coxeter transition on the corresponding matrix and 1.1.17 3) will cause a 2-Coxeter transition. Thus the theorem is an immediate consequence of Theorem 1.1.2 and Propositions 1.1.2 and 1.1.3.

2. Balanced Labeled Circle Diagrams

2.1 From matrices to tabloids

The matrix approach of Kassel et. al. has naturally brought us to the general notion of Balanced Labeled Circle Diagram introduced in [] and []. Although it will be good to keep in mind the mechanisms that produce the matrices $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ it will be more convenient to carry out all our combinatorial constructions and manipulations directly on these tabloids. Roughly speaking, these tabloids are obtained by filling the circles in the diagram of σ with the labels $1, 2, \ldots, l$ so that "k" is in the same position as "x_k" is in $P_w^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$.

To be precise, in view of Theorem 1.2.2, we have the following

Definition 2.1

Given a permutation σ of length l, here and after we associate to each word $w = a_1 a_2 \cdots a_l \in RED(\sigma)$ the tabloid $T(w)$ obtained by placing "k" in the "$\bigcirc$" that is in position (i, j) if and only if the transposition s_{a_k} interchanges i with σ_j.

Now it develops that these tabloids have a very curious characterization. To state it we need some notation and further definitions. To begin, it will be convenient to let "$CD(\sigma)$" denote the circle diagram of a permutation σ. If σ has length l then $CD(\sigma)$ has l circles and a filling of these circles with the labels $1, 2, \ldots, l$ will be called an "*injective*" labeling of $CD(\sigma)$ or briefly an "*injective tabloid*". The label in position (i, j) in the resulting tabloid T will be denoted T_{ij}. We shall of course use matrix convention to denote location and thus i increases as we go SOUTH and j increases as we go EAST. As we did for matrices, if T is an injective labeling of $CD(\sigma)$, we shall denote by $T(j_1, j_2, \ldots, j_k)$ the subdiagram of T that is contained in columns $j_1, j_2, \ldots, j_k$ and rows $\sigma_{j_1}, \sigma_{j_2}, \ldots, \sigma_{j_k}$. We shall also denote by $T_{rs}(j_1, j_2, \ldots, j_k)$ the entry that is in the r^{th} row and s^{th} column of $T(j_1, j_2, \ldots, j_k)$.

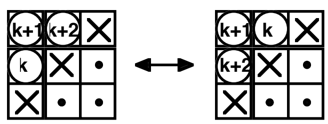
For a given cell $(i, j) \in CD(\sigma)$ the collection of cells that are directly EAST of (i, j) is called the "*arm*" of (i, j). Likewise the collection of cells that are directly SOUTH of (i, j) is called the "*leg*" of (i, j). The collection consisting of the cell (i, j) together with its arm and leg is usually referred to as the "*hook*" of (i, j), it will be denoted by "H_{ij}". A hook H_{ij} of an injective tableau T is said to be "*balanced*" if and only if the number of labels in the arm of (i,j) that are smaller than T_{ij} is equal to the number of labels in the leg that are bigger than T_{ij}. In particular we see that if the labels in H_{ij} are sorted in increasing order then placed back in H_{ij} starting from the bottom of the leg then NORTH up to (i,j) then finally EAST along the arm, T_{ij} will necessarily land right back in its cell. We say that T itself is "*balanced*" if all its hooks are balanced.

The notions of "*arm*", '*leg*", "*hook*" and "*balanced hook*" and "*balanced tabloid*" are easily extended to subdiagrams $T(j_1, j_2, \ldots, j_k)$. For instance we let the arm of $T_{r,s}(j_1, j_2, \ldots, j_k)$ be the collection of cells of $T(j_1, j_2, \ldots, j_k)$ that are EAST of $T_{r,s}(j_1, j_2, \ldots, j_k)$. The remaining notions are analogously defined. In particular, we let $H_{rs}(j_1, j_2, \ldots, j_k)$ denote the hook of $T_{r,s}(j_1, j_2, \ldots, j_k)$. To be precise, $H_{rs}(j_1, j_2, \ldots, j_k)$ consists of $T_{r,s}(j_1, j_2, \ldots, j_k)$ together with its arm and leg in $T(j_1, j_2, \ldots, j_k)$. Likewise we say that $T(j_1, j_2, \ldots, j_k)$ is balanced if all the hooks $H_{rs}(j_1, j_2, \ldots, j_k)$ are balanced.

It goes without saying that all the results we have established for the matrices $P^{\mathcal{J}}(x_1, x_2, \ldots, x_l)$ can be transfered to the tabloids $T(w)$. We shall use this fact here and after without necessarily spelling out in detail how this transfer should be carried out, since it only amounts to making the replacements

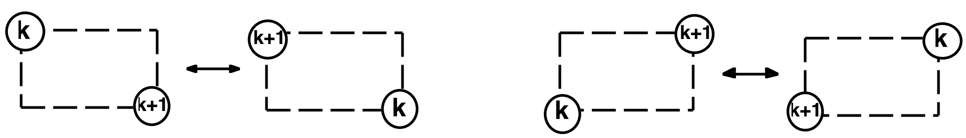
$$\text{“}x_k\text{”} \longrightarrow \text{“}\,k\,\text{”}\ , \qquad \text{“0”} \longrightarrow \text{“}\bullet\text{”}\ , \qquad \text{“1”} \longrightarrow \text{“}X\text{”}\ ,$$

In particular the 3-Coxeter and 2-Coxeter transitions of section 1.2 now become as indicated below. Namely, 3-Coxeter transitions are simply interchanges in T of 3×3-subdiagrams $T(j_1, j_2, j_3)$ of the form:



2.1.1

while 2-Coxeter transitions are substitutions of the form



2.1.2

In the same vein Theorem 1.2.3 may now be stated as

Theorem 2.1.1

For any two reduced words $w_1, w_2 \in RED(\sigma)$ we can find a sequence of Coxeter transitions which transform $T(w_1)$ into $T(w_2)$.

The notion of balanced tabloid arised quite early [] in the study of reduced words. The work of Kassel et. al. shows that it has a natural algebraic setting which beutifully explains its origin. We derive it here as a corollary of Theorem 2.1.1.

Proposition 2.1.1

The tabloids $T(w)$ are all balanced.

Proof

A view of the displays in 2.1.1 and 2.1.2 should make it clear that applying a 2 or 3-Coxeter transition on a balanced tabloid does not destroy balance. At any rate, note that in the case of the 3-Coxeter transition which goes from left to right in 2.1.1 we see that we are increasing by one the number of entries in the arm of $k+1$ that are less than $k+1$ but at the same time we are increasing by one the number of entries in the leg of $k+1$ that are larger than $k+1$. Going from right to left in 2.1.1 reverses this process and cannot affect balance of the hook of $k+1$. All the other hooks H_{ij} contain only $k, k+1$ or $k+1, k+2$ and their balance is trivially not affected by either of the two changes in 2.1.1. Likewise, the balance of a hook is not affected by any of the two transition in the first part of 2.1.2, for in this case no hook contains both k and $k+1$. As for the transitions in the second part of 2.1.2, note that if $T_{ij} \neq k, k+1$ then $T_{ij} > k$ if and only if $T_{ij} > k+1$ and the balance of H_{ij} cannot be affected by this transition. Similarly, if $T_{ij} = k$ or $T_{ij} = k+1$ then replacing k by $k+1$ or viceversa cannot affect the balance of H_{ij}.

To conclude, note that the tabloid $T(w_o)$ of any canonical factorization w_o is necessarily balanced since, by the way $T(w_o)$ is constructed (cf. Theorem 1.2.1), all the labels in the arm of a hook $H_{i,j}$ are larger than T_{ij} and all the labels in the leg are smaller. Now when w and w_o are reduced words of the same permutation, by Theorem 2.1.1, we can pass from $T(w_o)$ to $T(w)$ by a sequence of Coxeter transitions. Since when w_o is canonical $T(w_o)$ is balanced, $T(w)$ must be balanced as well since, as we have seen, all these transitions preserve balance.

We should note that Proposition 1.2.2 yields us an algorithm for constructing our tabloids $T(w)$ without resorting to multiplication diagrams. In fact, Proposition 1.2.2, converted to tabloids, may be restated as

Proposition 2.1.2

Let $w = a_1 a_2 \cdots a_k \in RED(\sigma)$, and let $\sigma_j < \sigma_{j+1}$ so that $w' = a_1 a_2 \cdots a_k\, j \in RED(\sigma \times s_j)$, then the transition

$$T(w) \longrightarrow T(w')$$

is simply obtained by interchanging columns j and $j+1$ of $T(w)$ and then changing the (σ_j, j)-entry of the resulting tabloid to "$k+1$".

This result as an immediate converse which may be stated as follows

Proposition 2.1.3

Let $w = a_1 a_2 \cdots a_k\, j \in RED(\sigma)$, and let $\sigma_j > \sigma_{j+1}$ so that $w' = a_1 a_2 \cdots a_k \in RED(\sigma \times s_j)$, then the transition

$$T(w) \longrightarrow T(w')$$

is simply obtained by interchanging columns j and $j+1$ of $T(w)$ and then changing the "$k+1$" to a "$\bullet$".

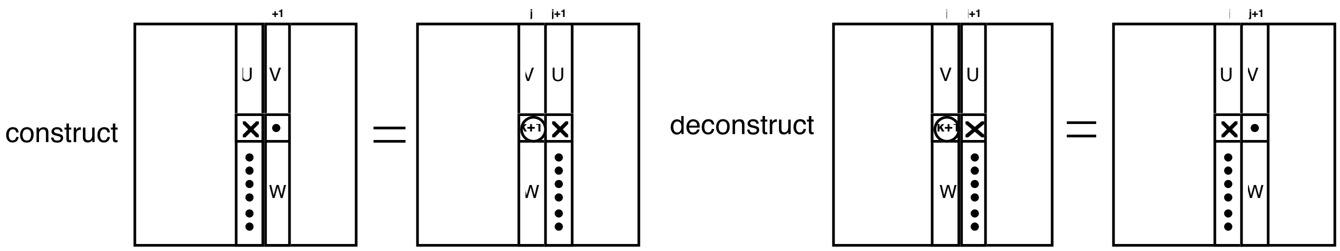
At this point it is good to have a visual image of these two transformations. For convenience let "**construct**" and "**deconstruct**" denote the transformations described in Propositions 1.1.2 and 2.1.3. More precisely when $w \in RED(\sigma)$ and $\sigma_j < \sigma_{j+1}$ then

$$\textbf{construct}\big[T(w), j\big] \;=\; T(w\, j)$$

and when $w = a_1 a_2 \cdots a_{k+1} \in RED(\sigma)$ then

$$\mathbf{deconstruct}\big[T(a_1 a_2 \cdots a_{k+1})\big] \;=\; \big(T(a_1 a_2 \cdots a_k),\, j\big)\,.$$

This given we can schematically represent Propositions 2.12 and 2.1.3 by the following displays.



Remark 2.1.1

We should note that to apply **construct** we need to give j and then k is the largest entry in $T(w)$. To apply **deconstruct** we locate the largest entry, say it is $k+1$ and it lies in the j^{th} column. This given we operate as indicated in the figure and return the resulting tabloid along with the index "j".

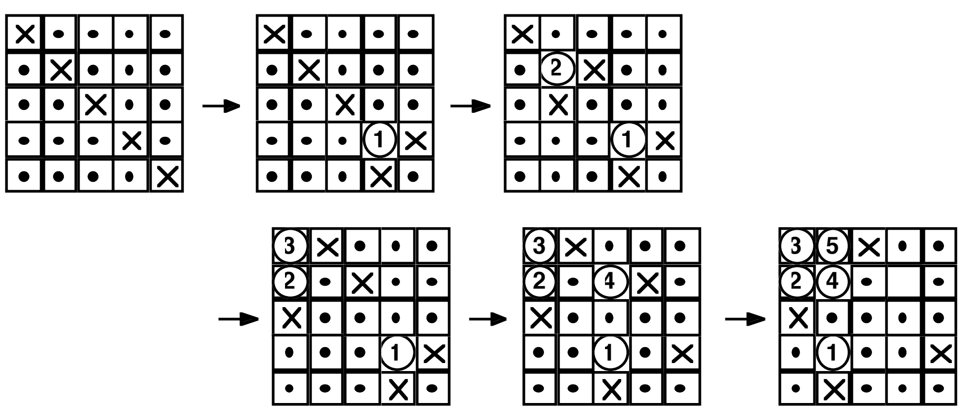
Now we see that to construct a tabloid $T(a_1 a_2 \cdots a_l)$ we only need to carry l applications of **construct**. More precisely, we recursively set

$$\mathbf{T}(\mathbf{a_1 a_2} \cdots \mathbf{a_{k+1}}) \;=\; \mathbf{construct}\big[\mathbf{T}(\mathbf{a_1 a_2} \cdots \mathbf{a_k}), \mathbf{a_{k+1}}\big] \qquad (\text{for } k = 1, 2, \ldots, l-1)$$

with the initial step

$$\mathbf{T}(\mathbf{a_1}) \;=\; \mathbf{construct}\big[\mathbf{T_o}\big]$$

where T_o is the tabloid that corresponds to the identity permutation. In the following display we have carried out this algorithm for $w = 42132$.



For a moment let us say that a injective labeling T of the circle diagram $CD(\sigma)$ is "*constructible*" if and only if $T = T(a_1 a_2 \cdots a_l)$ for some $a_1 a_2 \cdots a_l \in RED(\sigma)$.

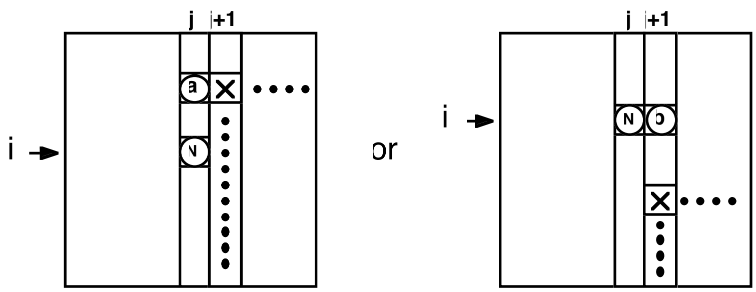
We have the following remarkable fact.

Theorem 2.1.2

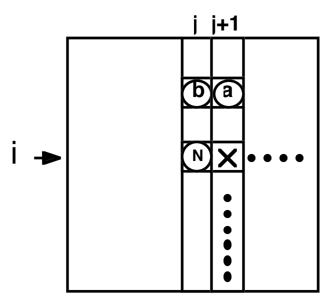
An injective labeling of the circle diagram of a permutation is constructible if and only if it is balanced.

Proof

In view of Proposition 2.1.1, we need only prove that every balanced tabloid is constructible. Let then T be a balanced labelling of the circle diagram of σ and let N be the largest label in T. Suppose further that $N = T_{ij}$. We claim that in position $(i, j+1)$ there necessarily is an X. To see this, note that if this were not so then the j^{th} and $j+1^{st}$ columns of T would have one of the following forms:



Indeed, if the X in column $j+1$ were above the th row then immediately to the left of it there would have to be a circle because there is no "X" to kill that cell from the left or above. Now the label in that circle is necessarily a number $a < N$ but that would cause the hook of a to be unbalanced since there is a label bigger than a SOUTH of a and no label less than a EAST. In fact no label at all EAST of a because of that adjacent X. This eliminates the first alternative. In case the X in column $j+1$ is below the i^{th} row then there would have to be a circle in column $j+1$ immediately to the right of N because there is no "X" to kill that cell from the left or from above. Now again in that circle there would have be a label $b < N$, but then the hook of N is unbalanced be cause there is a label smaller than N EAST and no label bigger than N SOUTH. This eliminates the second possibility. This forces the j^{th} and $j+1^{st}$ columns to be of the following form

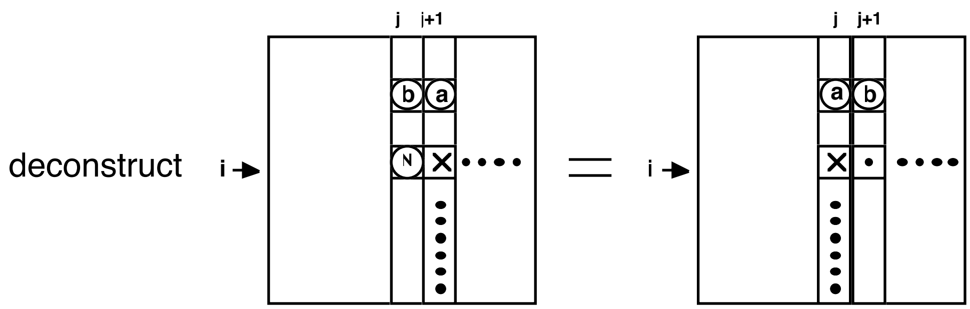


2.1.3

where we claim that every label b above N, in the j^{th} column, has necessarily an adjacent label $a < b$ in the $j+1^{st}$ column. Clearly, there must be a circle adjacent to b in the $j+1^{st}$ column because there is no "X" to kill that cell from the left or from above. To show that in that circle there is a label a less than b we proceed by contraddiction. Suppose that the situation is as indicated in 2.1.3 with $a > b$, and that pair is the <u>lowest</u> we can find. Let then p be the number of labels, SOUTH of b, that are larger than b. We have $p \geq 1$ because $N > b$. But then, since T is balanced, there must also be p labels $b_1, b_2, \ldots, b_p$ all less than b in the arm of b.

Now, since $a > b$ all these labels are less than a as well. But then again, since T is balanced, there must be at least p labels $u_1, u_2, \ldots, u_p$ larger than a in the leg of a. However all these labels must fall in circles of column $j+1$ that are between the a and the X. Moreover, the presence of these circles in column $j+1$ forces circles adjacent to them in column j. Let $w_1, w_2, \ldots, w_p$ be the labels that fall in these circles, (indexed so that w_r is to the left of u_r). Since we chose b and a to form the lowest pair $b < a$, we must have $w_r > u_r > a > b$ (for $r = 1, 2, \ldots, p$). In summary, these two columns would the be as depicted in the adjacent figure. But this cannot be since we now see $p+1$ labels greater than b in the leg of b, contrary to our initial choice of p. We have now proved that T is the form given in 2.1.3 where every pair of adjacent circles above the pair N, X contain labels b, a with $b > a$. We claim that if we apply **deconstruct** to T the resulting tabloid T' will be again a balanced injective labelling of $CD(\sigma)$. Indeed a look at the picture below should make it clear that the only hooks whose collections of labels have been affected in a significant way are those of a and b. Now a only gains a label greater than it to the right, this does not affect its rank among the labels in his hook, so its hook remains balanced. As for b we see that it looses $N > b$ in its leg but at the same time it looses $a < b$ in its arm. These losses compensate each other and thus leave the hook of b still balanced.

$$
\begin{array}{ccc}
j & & j+1 \\
\end{array}
\begin{bmatrix}
\cdot & & \cdot \\
\cdot & & \cdot \\
\cdot & & \cdot \\
b & < & a \\
\cdot & & \cdot \\
\cdot & & \cdot \\
w_1 & > & u_1 \\
\cdot & & \cdot \\
w_2 & > & u_2 \\
\cdot & & \cdot \\
\cdot & & \cdot \\
\cdot & & \cdot \\
w_p & > & u_p \\
\cdot & & \cdot \\
\cdot & & \cdot \\
N & & X \\
\cdot & & \bullet \\
\cdot & & \bullet \\
\cdot & & \bullet \\
\cdot & & \bullet \\
\cdot & & \bullet
\end{bmatrix}
$$



We can see now how the proof can be completed. To begin the result is trivially true for the circle diagram of the identity since there are no circles at all to fill. So we assume by induction on the number of circles, that all balanced labelings of $CD(\sigma)$ that have less circles than T are constructible. Now we see from the figure above that T' is, in fact, a labelling of the circle diagram of the permutation $\sigma \times s_j$. The inductive hypothesis gives that T' is constructible. This given, we must have that $T' = T(a_1, a_2, \cdots a_{l-1})$ with $a_1, a_2, \cdots a_{l-1} \in RED(\sigma \times s_j)$ and a fortiori $a_1, a_2, \cdots a_{l-1}\ j \in RED(\sigma)$. Since

$$\mathbf{construct}\,(\mathbf{T}, \mathbf{j}) \quad = \quad \mathbf{T}'$$

We deduce that

$$T \quad = \quad T(a_1, a_2, \cdots a_l) \qquad \text{(with} \quad a_l = j)$$

This shows that T is constructible, completing the induction and the proof.

It develops that constructibility, (and now in particular balance) forces a whole family of restrictions on the labeling.

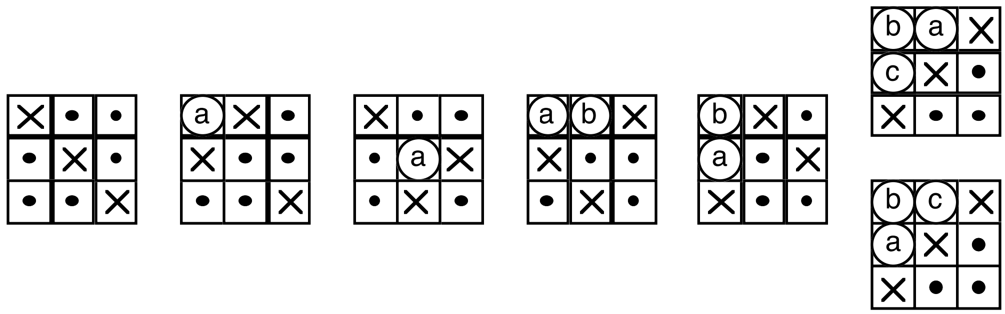
Definition 2.1.2

Let T be an injective labelling of the circle diagram of a permutation. We shall say that T is "k-balanced" if and only if all of its $k \times k$-subtabloids are balanced.

Now we have the following remarkable result.

Proposition 2.1.4

Every constructible tabloid is 3-balanced. In other words each of its 3×3 subtabloids must have one of the following forms with $a < b < c$:



2.1.4

Proof

Let $T = T(w)$ with $w = a_1 a_2 \cdots a_l \in RED(\sigma)$ and let $\mathcal{M}(a_1 a_2 \cdots a_l)$ be the multiplication diagram for the matrix

$$P_w(x_1, x_2, \ldots, x_l) \;=\; P_{a_1}(x_1) P_{a_2}(x_2) \cdots P_{a_l}(x_l)\,.$$

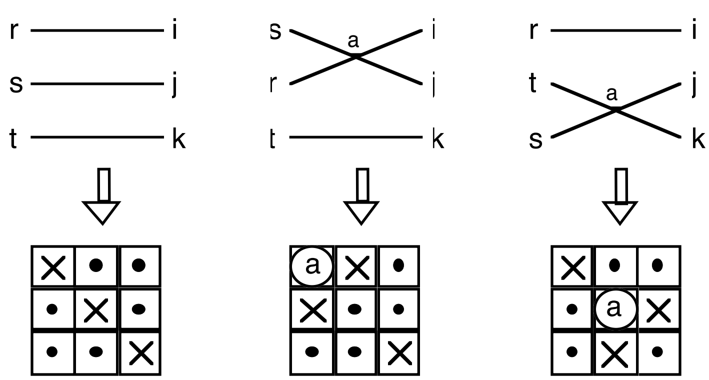
Let $1 \leq i < j < k \leq l$ be given indices and let

$$r = \sigma_i \quad , \quad s = \sigma_j \quad , \quad t = \sigma_k \quad .$$

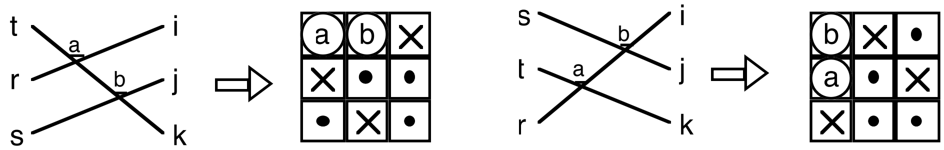
Now there are 6 possibilities.

$$r < s < t \quad , \quad s < r < t \quad , \quad r < t < s \quad , \quad t < r < s \quad , \quad s < t < r \quad , \quad t < s < r \quad ,$$

In the first case the r, s and t-lines do not cross in $\mathcal{M}(a_1 a_2 \cdots a_l)$. In the second case the s-line and r-line cross and in the third case it is the the t-line and s-line that cross. Assuming that these crossings occur at time a, we have schematically represented below, what these conditions imply on $\mathcal{M}(a_1 a_2 \cdots a_l)$ and the subdiagram $T(i, j, k)$:

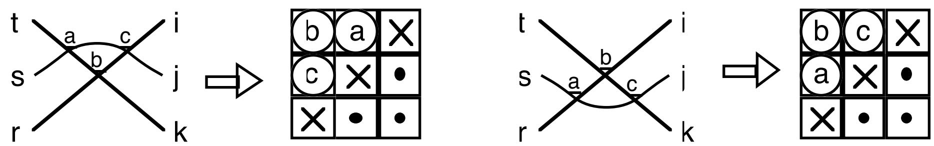


In the fourth and fifth case there are two crossings. Assuming that the first crossing occurs at time a and the second at time $b > a$, the fact that any two lines cross only once forces the line diagrams and implications depicted below:



In fact, for instance, when when $t < r < s$ we clearly see from the figure that the s and r lines must cross an even number of times and thus in a line diagram they can only cross 0 times. This given the t and r lines will necessarily be the first pair to cross.

Finally, when $t > s > r$, the t and r lines must cross, and if the time they cross is b, there still remains two possibilities. Indeed the geometry of these line diagrams requires the s line to go either over or under the b-crossing (see figure below). In the first case the t and s line crossing occurs first and the s and r lines cross last. In the second case the order is reversed. This accounts for the diagrams and implications depicted below.



This establishes our result.

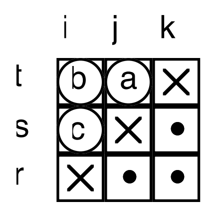
Now it develops that Proposition 2.1.4 can be reversed.

Proposition 2.1.5

Every 3-balanced injective tabloid is balanced, therefore constructible.

Proof

Let T be given and 3-balanced injective labeling of $CD(\sigma)$ and let $b = T_{ti}$. To show that the hook H_{ti} is balanced we need to show that there is a one to one correspondence between the labels "a" EAST of b that are less than b and the labels"c", SOUTH of b, that are larger than b. Now, this correspondence is simply obtained from the highest pattern in 2.1.4. To be precise, let $a < b$ be in position (t, j), with $j > i$ and let the "X" in row t be in column $k > j$. This given, from the list in 2.1.4 we deduce that the subdiagram $T(i, j, k)$ can only be of the form



with $t = \sigma_k$, $s = \sigma_j$ and $r = \sigma_i$. This shows that if $b = T_{ti}$ then the label $c > b$ in the leg of b that corresponds to a label $a < b$ in position (t, j) will be found in position (σ_j, i). This completes the argument.

What now follows by putting together all the results of this section is truly remarkable.

Theorem 2.1.3

For an injective labeling T of the circle diagram of a permutation σ of length l the following conditions are equivelent

(i) T is balanced,
(ii) T is constructible,
(iii) T is 3-balanced,
(iv) T is k-balanced for some $k = 4, 5, \ldots, l$,
(v) For some $k = 4, 5, \ldots, l$ all the $k \times k$ subtabloids of T are constructible.

Proof

We have the following sequence of implications

$$(i) \longrightarrow (ii) \longrightarrow (iii) \longrightarrow (iv) \longrightarrow (v) \longrightarrow (iii) \longrightarrow (i)$$

Indeed if T is balanced then it is constructible by Theorem 2.1.2. If it is constructible it is 3-balanced by Proposition 2.1.4. If it is 3-balanced all the $k \times k$ subtabloids are necessarily also balanced by Proposition 2.1.5 and so they are salso constructible by Theorem 2.1.2. But then all the 3×3 subtabloids will be balanced by Proposition 2.1.4 and then Proposition 2.1.5 yields that T must be balanced.

Remark 2.1.2

We should note that saying "*all the $k \times k$ subtabloids of T are constructible*" is an abuse of terminology. What we really should say that if we take a $k \times k$ subtabloid $T(j_1, j_2, \ldots, j_k)$ with labels $a_1 < a_2 < \cdots < a_m$ and respectively, replace these labels by $1, 2, \ldots, m$ the resulting tabloid $T'(j_1, j_2, \ldots, j_k)$ is constructible. For later purposes it will be good to formalize this operation, refering to it as "*downscaling*" and set

$$T'(j_1, j_2, \ldots, j_k) \;=\; \textbf{downscale}\left(T(j_1, j_2, \ldots, j_k)\right). \qquad 2.1.5$$

It follows then from Theorem 2.1.3 that if we want to have the list of all possible tabloids T' that may be obtained by downscaling a subtabloid $T(j_1, j_2, \ldots, j_k)$ we simply list all tabloids T_w corresponding to reduced decompositions of permutations in S_k.

We should also keep in mind that we denote by $\mathcal{M}(a_1, a_2, \cdots, a_l)$ the diagram corresponding to the product

$$P_{a_1}(x_1) P_{a_2}(x_2) \cdots P_{a_l}(x_l)$$

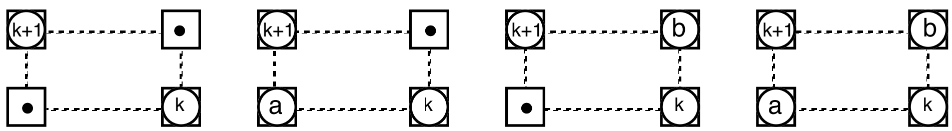
We will also informally refer to $\mathcal{M}(a_1, a_2, \cdots, a_l)$ as the "*line diagram for w*".

2.2 Descents and Kevin Kadel's ZIGZAGs

We define the "*Descent Set*" of a word $w = a_1 a_2 \cdots a_l$ and denote it "$D(w)$" the set

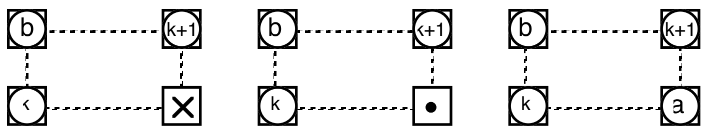
$$D(w) \;=\; \{1 \le k < l \,:\, a_k > a_{k+1}\} \qquad 2.2.1$$

Kevin Kadel discovered a remarkably simple way to recover $D(w)$ directly from the tabloid $T(w)$. To state it we need some notation. Note first that, if the labels k and $k+1$ are not in the same row or column of the tabloid $T(w)$, then the 2×2 subtabloid of $T(w)$ containing the labels k, $k+1$ may have one of the following forms.



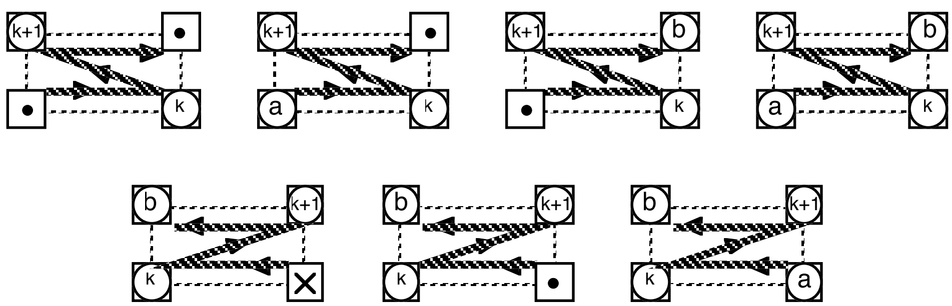
2.2.2

or



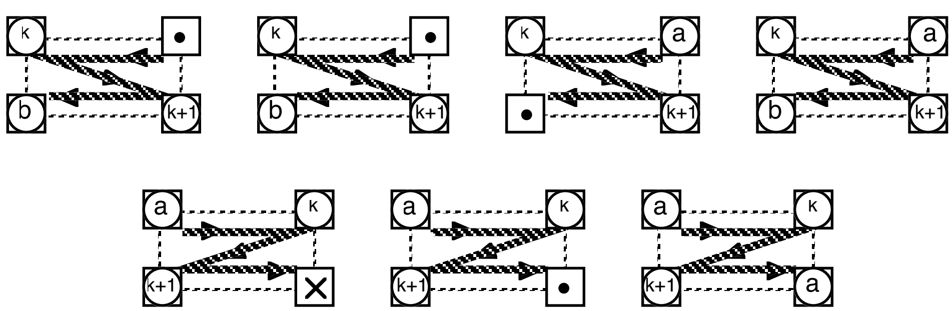
2.2.3

and, of course, also those that are be obtained from them by interchanging k and $k+1$. Kadell [] associates to each of these subtabloids a "*ZIGZAG*" path, whose midcorners are labelled by $k, k+1$, oriented so that k comes before $k+1$. The following display depicts the ZIGZAGs associated to the tabloids in 2.2.2 and 2.2.3.



2.2.4

The display below shows the ZIGZAGs obtained when we interchange k and $k+1$ in 2.2.2 and 2.2.3.



2.2.5

This given, Kadell's result may be stated as follows

Theorem 2.2.1

Let σ be of length l and $w = a_1 a_2 \cdots a_l \in RED(\sigma)$, then a given k is in the descent set $D(w)$ if and only if one of the following three conditions are satified

(1) *k and $k+1$ are in the same column of $T(w)$.*
(2) *k and $k+1$ are not in the same row or column of $T(w)$, k is in a lower row than $k+1$ and the corners of their ZIZAG contain no other labels.*
(3) *k and $k+1$ are not in the same row or column of $T(w)$ and the labels encountered in their ZIGZAG are in increasing order.*

Proof

We should note that (3) simply means that the ZIGZAG of k and $k+1$ is given by one of the patterns in in 2.2.4 and 2.2.5 with $a < k$ and $b > k+1$. We shall first establish the result under the assumption that $k+1$ is the largest label. More precisely we work with $\sigma \to \sigma^{(k+1)} = s_{a_1} s_{a_2} \cdots s_{a_{k+1}}$ and $w \to a_1 a_2 \cdots a_{k+1}$. This given, letting $r = a_k$ and $s = a_{k+1}$, the last two columns of the line diagram $\mathcal{M}(w^{(k+1)}$ can be schematically represented by one of the four cases depicted below

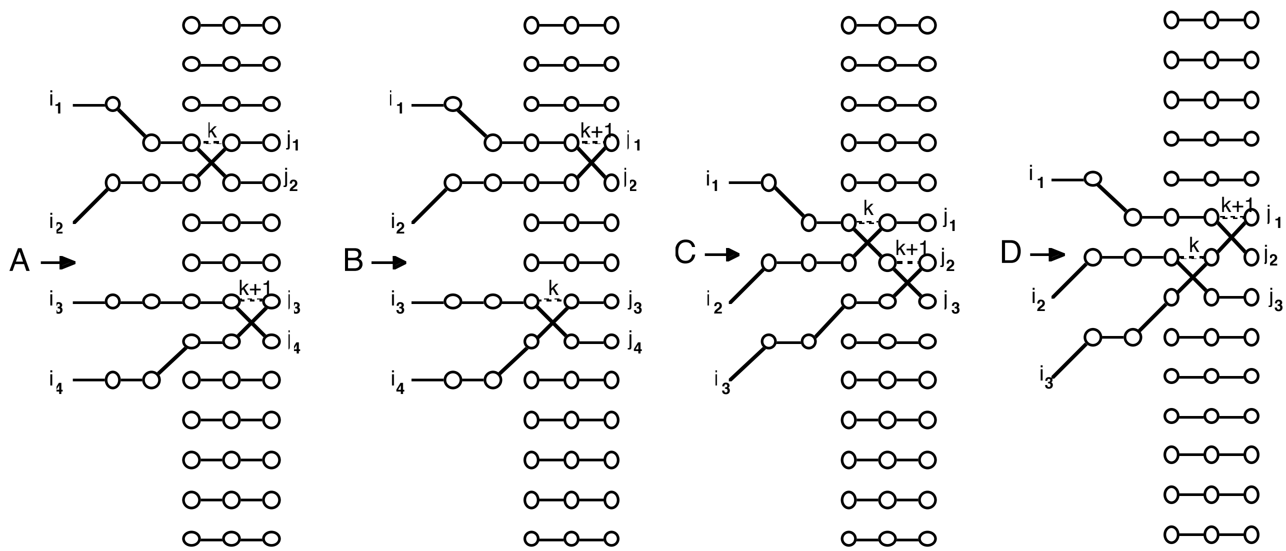


Fig 2.2.5

Cases A and B occur when $|r-s| > 1$, cases C and D when $|r-s| = 1$. Moreover

(ac) Cases A and C occur when k **is not** a descent, that is we have $a_k < a_{k+1}$ (i.e. $r < s$) .

(bc) Cases B and D occur when k **is** a descent, that is we have $a_k > a_{k+1}$ (i.e. $r > s$) .

The indices i_1, i_2, i_3, i_4 and j_1, j_2, j_3, j_4 are determined as follows

(1) In cases A:

$$\begin{aligned} &i_1=\sigma^{(k+1)}_{j_2},\ \ i_2=\sigma^{(k+1)}_{j_1},\ \ i_3=\sigma^{(k+1)}_{j_4},\ \ i_4=\sigma^{(k+1)}_{j_3},\\ &j_1=r\ \ ,\qquad j_2=r+1\ \ ,\ \ j_3=s\ \ ,\qquad j_4=s+1 \end{aligned}\tag{2.2.6}$$

(2) In cases B:

$$\begin{aligned} &i_1=\sigma^{(k+1)}_{j_2},\ \ i_2=\sigma^{(k+1)}_{j_1},\ \ i_3=\sigma^{(k+1)}_{j_4},\ \ i_4=\sigma^{(k+1)}_{j_3},\\ &j_1=s\ \ ,\qquad j_2=s+1\ \ ,\ \ j_3=r\ \ ,\qquad j_4=r+1 \end{aligned}\tag{2.2.6}$$

(3) In case C we have $s=r+2$ and

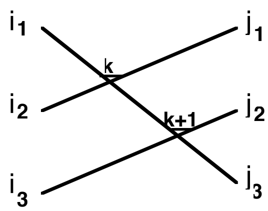
$$\begin{aligned} &i_1=\sigma^{(k+1)}_{j_3},\ \ i_2=\sigma^{(k+1)}_{j_1},\ \ i_3=\sigma^{(k+1)}_{j_2},\\ &j_1=r\ \ ,\qquad j_2=r+1\ \ ,\ \ j_3=r+2\ \ . \end{aligned}\tag{2.2.7}$$

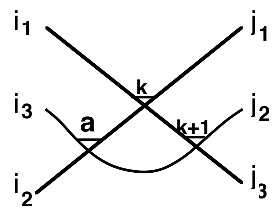
(4) In case D we have

$$\begin{aligned} &i_1=\sigma^{(k+1)}_{j_2},\ \ i_2=\sigma^{(k+1)}_{j_3},\ \ i_3=\sigma^{(k+1)}_{j_1},\\ &j_1=r\ \ ,\qquad j_2=r+1\ \ ,\ \ j_3=s\ \ . \end{aligned}\tag{2.2.8}$$

It develops that, in case C the labels k and $k+1$ are in the same row of the tabloid $T(a_1a_2\cdots a_{k+1})$ and in case D they are in the same column. This is in complete agreement with with assertion (1) of the Theorem.

To prove this note first that, in case C, if w continue the i_1, i_2 and i_3 lines all the way to the beginning of the diagram $\mathcal{M}(a_1a_2\cdots a_{k+1})$ the i_1-line cannot intersect either of the i_2 and i_3 lines but the i_2 and i_3 lines can intersect. Now using the assignements in 2.2.7 we can schematically represent these two possibilities by the following diagrams



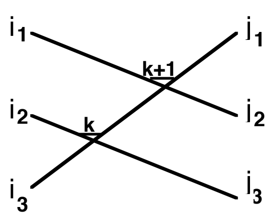


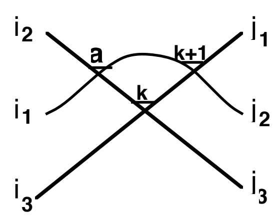
form which we derive the following two possibile forms for the the subtabloid $T_{w^{(k+1)}}(j_1,j_2,j_3)$ in case C :

	j_1	j_2	j_3
i_1	k	k+1	×
i_2	×	•	•
i_3	•	×	•

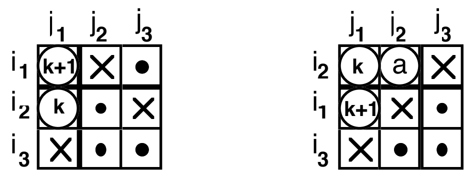
	j_1	j_2	j_3
i_1	k	k+1	×
i_2	a	×	•
i_3	×	•	•

Likewise from the assignments in 2.28 we derive that case D the portion of $\mathcal{M}(a_1a_2\cdots a_{k+1})$ consisting of the i_1, i_2 and i_3 lines can be schematically represented by one of the following two diagrams





consequently, in this case we get the following two possibile forms for the the subtabloid $T_{w^{(k+1)}}(j_1, j_2, j_3)$:



We show next that in the Cases A and B the labels k and $k+1$ are at the corners of a rectangle with the cooresponding ZIGZAG meeting its labels in increasing order in Case B and in disorder in Case A. Note that since the word corresponding to Case A can be obtained from the word corresponding to Case B by interchanging the last two letters, we derive that the tabloid $T(a_1a_2\cdots a_{k+1})$ corresponding to case A can be obtained from that corresponding to Case B by a 2-Coxeter transition. This implies that the ZIGZAGs occurring in case A can be obtained from those occurring in case B by interchanging k and $k+1$. Therefore it will be sufficient to show that, in Case B the ZIGZAGs meet their labels in increasing order. It will then follow that, in Case B, the ZIGZAGs meet their labels in disorder as asserted.

We can deal with case B as we did for Cases C and D. We start by noting that, as we follow the i_1, i_2, i_3 and i_4 lines from the last three columns of the diagram $\mathcal{M}(a_1a_2\cdots a_{k+1})$, all the way back to the beginning, there cannot be any further intersections between the i_1 and i_2 lines nor between the i_3 and i_4 lines. This implies that the indices i_1, i_2, i_3, i_4 may be governed only by the following six sets of inequalities:

$$
\begin{array}{lll}
1)\ i_1 < i_2 < i_3 < i_4 & 2)\ i_1 < i_3 < i_2 < i_4 & 3)\ i_1 < i_3 < i_4 < i_2 \\
4)\ i_3 < i_1 < i_2 < i_4 & 5)\ i_3 < i_1 < i_4 < i_2 & 6)\ i_3 < i_4 < i_1 < i_2
\end{array}
$$

In the two following displays we have schematically represented each ensuing diagram and the corresponding form of the subtabloid $T_{w^{(k+1}}(j_1, j_2, j_3, j_4)$. Here we have omitted filling the circles that are not corners of the ZIGZAG of k and $k+1$. We should note that the label "a", of course, will always be less that k since the corresponding intersection occurs before time k.

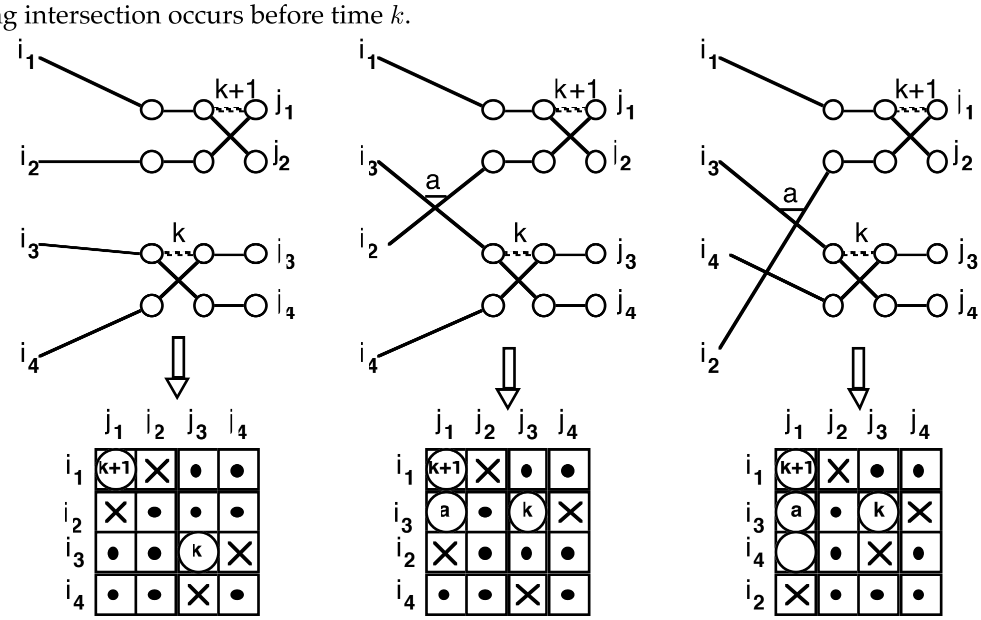


Fig 2.2.9

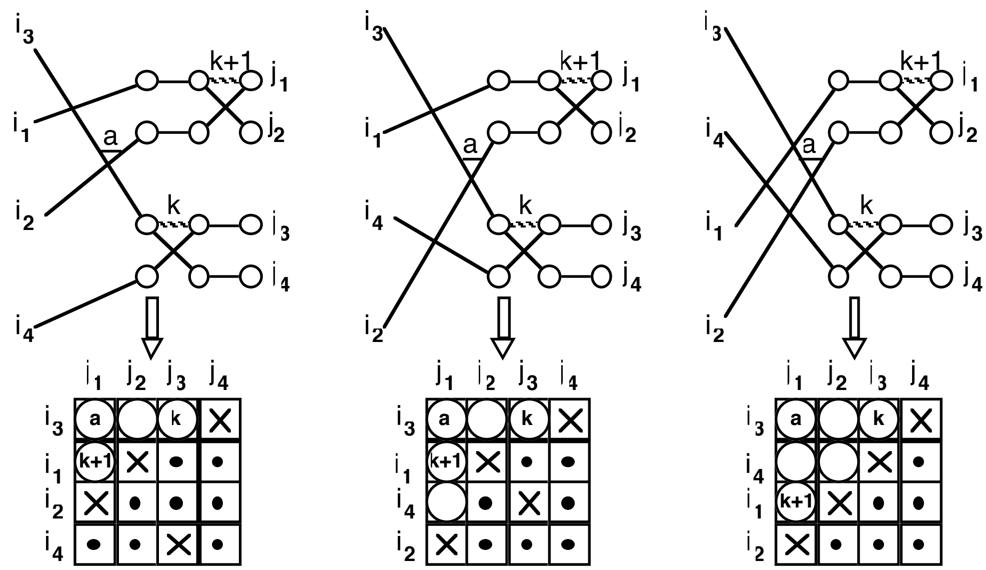
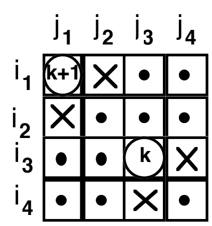


Fig. 2.2.10

We have thus established the result in the case that $k+1$ is the highest label. To complete the proof we need only check what happens as we continue the diagram $T(a_1a_2\cdots a_{k+1})$ so as to obtain the final diagram $T(a_1a_2\cdots a_{k+1}a_{k+2}\cdots a_l)$. To begin with we note that in applying "**construct**" a "•" can never be changed to a labelled "O" if it lies below an "X". This means that the first subtabloid $T_{w(k+1)}(j_1,j_2,j_3,j_4)$ in 2.2.9 namely



will never acquire a labelled circle below $k+1$ in the row of k. Moreover if it acquires a circle in the row of $k+1$ above k it will necessarily be with a label $b>k+1$ yielding a ZIGZAG with increasing labels. Likewise, in all remaining cases of Figures 2.2.9 and 2.2.10, any labelled circle added to a further corner of the ZIGZAG will also come with a label $b>k+1$. Finally it is easily seen that the transformations which $T_{w(m)}(j_1,j_2,j_3,j_4)$ undergoes as m increases to l under successive applications of "**construct**" cannot change an ordered ZIGZAG to a disordered one for the simple reason that the label $a<k$ will always remain before k and the label $b>k+1$ will always remain after $k+1$ in the ZIGZAG ordering. This completes our argument.

2.3 Special Circle Diagrams

In this section we introduce several important classes of permutations and derive some useful properties of their circle diagrams. Before we can proceed we need to make a few definitions. To begin, for a given $\sigma \in S_n$ we set for $1 \leq i \leq n-1$

$$C_i(\sigma) \;=\; \big\{\, j > i \,:\, \sigma_j < \sigma_i \,\big\}\,. \qquad 2.3.1$$

The sequence of subsets

$$C(\sigma) \;=\; [C_1(\sigma), C_2(\sigma), \ldots, C_{n-1}(\sigma)]\,, \qquad 2.3.2$$

will be referred to as the "*code sequence*" of σ. Setting

$$c_i(\sigma) \;=\; \#\big\{\, j > i \,:\, \sigma_j < \sigma_i \,\big\} \;=\; |C_i(\sigma)|\,, \qquad 2.3.3$$

the vector

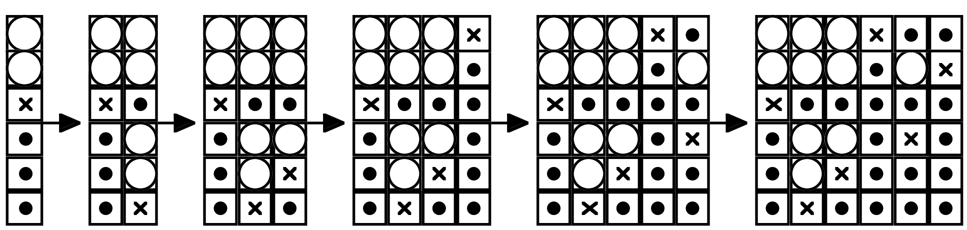
$$c(\sigma) \;=\; \big(c_1(\sigma), c_2(\sigma), \ldots, c_{n-1}(\sigma)\big)$$

will be called the "*code*" of the permutation σ.

Note that we have

$$c_i(\sigma) \leq n-i \qquad \big(\text{for } i = 1, \ldots, n\big)\;, \qquad 2.3.4$$

this is because there are i "X" 's in the first i columns of $CD(\sigma)$ that leaves at most $n-i$ cells in the i^{th} column where we can put a circle. It is also easily seen that every vector $c = (c_1, c_2, \ldots, c_n)$ with non-negative integer components satisfying the inequalities in 2.3.4 is the code of a permutation $\sigma \in S_n$. Indeed, the circle diagram $CD(\sigma)$, and therefore σ itself are easily reconstructed form $c(\sigma)$. We start by placing an "O" ' in each of the first c_1 cells of the first column of $CD(\sigma)$ followed by an "X" in the $c_1 + 1^{st}$ cell. Then, having placed all the "X" 's, the "O" 's and the "•" 's in the first $i-1$ columns, we fill the i^{th} column by first placing the "•" 's in each cell that is killed by an "X" to its WEST, then place the "O" 's in the first c_i available cells followed by an "X" in the next available cell. Of course here "available cell" means a cell that has not been killed by a previous "X". Note that after we filled the $n-1^{st}$ column, the n^{th} column will automatically get an "X" in the only remaining available cell. In the display below we illustrate this construction process when the given code is $c = (2, 4, 3, 0, 1, 0)$ yielding the permutation $\sigma = 365142$.



Let γ be a two-line array

$$\gamma \;=\; \begin{bmatrix} j_1 & j_2 & j_3 & \cdots & j_k \\ a_1 & a_2 & a_3 & \cdots & a_k \end{bmatrix}$$

with $j_1 < j_2 < j_3 \cdots < j_k$ and $a_1, a_2, a_3, \ldots, a_k$ distinct integers. For a given $i = 1, 2, \ldots, k$ let $r_i(\gamma)$ denote the "*rank*" of a_i in the set $\{a_1, a_2, a_3, \ldots, a_k\}$. That is we set $r_i(\gamma) = m+1$ if and only if precisely m of $a_1, a_2, a_3, \ldots, a_k$ are less than a_i. This given, we shall say that "γ *downscales*" to the permutation

$$r(\gamma) \;=\; \begin{bmatrix} 1 & 2 & 3 & \cdots & k \\ r_1(\gamma) & r_2(\gamma) & r_3(\gamma) & \cdots & r_k(\gamma) \end{bmatrix} \;. \qquad 2.3.5$$

Let $\sigma \in S_n$ and let $\theta \in S_k$ for some $2 \leq k \leq n$, we shall say that σ is "*θ-avoiding*" if we cannot find indices $1 \leq j_1 < j_2 < \cdots < j_k \leq n$ such that the two-line array

$$\gamma = \begin{bmatrix} j_1 & j_2 & j_3 & \cdots & j_k \\ \sigma_{j_1} & \sigma_{j_2} & \sigma_{j_3} & \cdots & \sigma_{j_k} \end{bmatrix}$$

downscales to θ.

Remark 2.3.1

It is not difficult to see that a permutation σ is θ-avoiding if and only if there are no subdiagrams of the circle diagram of σ which are identical to the circle diagram of θ.

We now have the following remarkable result

Theorem 2.3.1

If a permutation σ is 321-avoiding then

(i) When we remove from $CD(\sigma)$ all the rows and columns that contain no circles, the circles in the resulting diagram fill the cells of a French skew Ferrers diagram D.

(ii) For every $w \in RED(\sigma)$ the balanced filling T_w of $CD(\sigma)$ can be obtained from a corresponding standard filling τ_w of D.

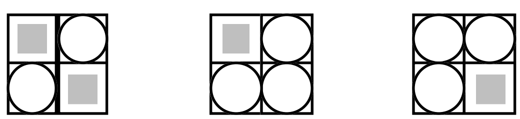
(iii) The descent sets of w and τ_w are identical.

Proof

A French skew diagram D is characterized by the following property

$$(i_1, j_1), (i_2, j_2) \in D \quad \text{with} \quad i_1 \leq i_2 \;\&\; j_1 \leq j_2 \quad \longrightarrow \quad (i_1, j_2), (i_2, j_1) \in D$$

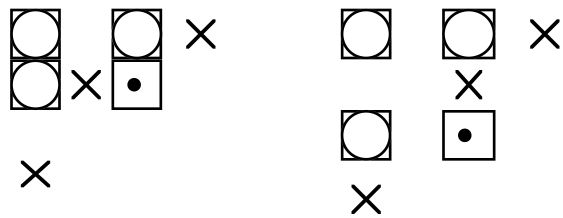
Thus to prove the first assertion we need only show that no 2×2 subdiagram of $CD(\sigma)$ can have any of the following forms



where a shade in a cell signifies absence of a circle. To begin with, it is easily seen that the first two cases can never occur for a 2×2 subdiagram of a circle diagram. To eliminate the third case let it be possible that the 2×2 subdiagram of $CD(\sigma)$ contained in rows i_1, i_2 and columns j_1, j_2 has any of the two forms below

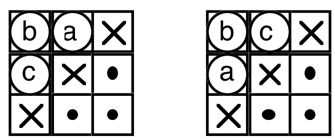
2.3.6

Now in the first case the "$\bullet$" must be due to an "X" either to the left or above. However, this "X" together with the "X"'s in column j_1 and row i_1 yield us one of the two configurations below

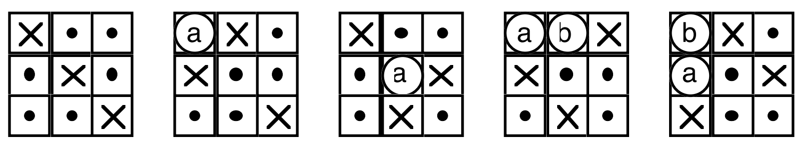


which yield a 321-pattern that σ is supposed to avoid. Likewise for the second case in 2.3.6 the "X"'s in column j_1 and row i_1 yield us again a 321 pattern. In either case we reach a contraddiction. This proves (i).

As for *(ii)* note that, since σ is 321-avoiding, the two tabloids below



cannot occur as 3×3 subtabloids of T_w. Thus from Propostion 2.1.4 we derive that the only remaining possibilities for a 3×3 subtabloid of T_w are



with $a < b$. This means that for any pair of labels appearing in the same row of T_w the one to the left is smaller that the one to the right and for any pair of labels that are in the same column the one below is smaller than the one above. This shows that this filling can be obtained from (or gives rise to) a standard filling τ_w of D.

To prove *(iii)* recall that k is called a "*descent*" of a french standard tableau if and only if $k+1$ is NORTH-WEST of k. This given, we see that if k and $k+1$ are in the same row in T_w then $k+1$ is to the right k and therefore k is not in the descent set of τ_w. If k and $k+1$ are in the same column of T_w then $k+1$ is above k and therefore k is in the descent set of τ_w. Finally, if k and $k+1$ are not in the same row or column and there are no other labels in the in the ZIGZAG of k and $k+1$ then from (2) of Theorem 2.2.1 we get that we have a descent at k for w if and only if $k+1$ is NORTHWEST of k in T_w. This makes k also a descent of τ_w. Likewise, if the ZIGZAG of k and $k+1$ as some other labels then from (3) of Theorem 2.2.1 we get that k is a descent of w if and only if the labels in the ZIGZAG of k and $k+1$ are in increasing order. But that again can happen if and only if $k+1$ is NORTHWEST of k in T_w. In any case we see that the assertion in *(iii)* is an immediate consequence of Kadel's Theorem 2.2.1. This completes our proof.

Definition 2.3.1

The decreasing rearrangement of the code of a permutation σ (with all the zero's omitted) will be here and after called the "shape of σ" and will be denoted $\lambda(\sigma)$.

The following is an important property of the shape.

Proposition 2.3.1

For any permutation σ we have

$$\lambda(\sigma) \leq \lambda'(\sigma^{-1}) \tag{2.3.7}$$

where "$\leq$" represents the dominance partial order. Moreover, equality here holds if and only if the code sequence $C(\sigma)$ is totally ordered by set inclusion.

Proof

Let $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ and let $M = \|m_{i,j}\|_{i,j=1}^{n}$ with

$$m_{i,j} \;=\; \begin{cases} 1 & \text{if there is a circle in } CD(\sigma) \text{ in position } (i,j) \\ 0 & \text{otherwise} \end{cases}$$

It is easily derived from the definition in 2.3.3 that the column sums of M are $c_1(\sigma), c_2(\sigma), \ldots, c_n(\sigma)$ and the row sums are $c_1(\sigma^{-1}), c_2(\sigma^{-1}), \ldots, c_n(\sigma^{-1})$. Let $j_1, j_2, \ldots, j_n$ be a permutation that rearranges $c_1(\sigma), c_2(\sigma), \ldots, c_n(\sigma)$ in decreasing order so that

$$c_{j_1}(\sigma), c_{j_2}(\sigma), \ldots, c_{j_n}(\sigma)$$

except for some terminal zeros gives $\lambda(\sigma)$. Clearly, the matrix

$$M' \;=\; \begin{pmatrix} m_{1,j_1} & m_{1,j_2} & \ldots & , m_{1,j_n} \\ m_{2,j_1} & m_{2,j_2} & \ldots & , m_{2,j_n}) \\ \vdots & \vdots & \vdots & \vdots \\ m_{n,j_1} & m_{n,j_2} & \ldots & , m_{n,j_n} \end{pmatrix}$$

has the same row sums as M and moreover, for every $k = 1, 2, \ldots, n$, the number of 1's in the first k columns of M' is given by

$$c_{j_1}(\sigma) + c_{j_2}(\sigma) + \cdots + c_{j_k}(\sigma)$$

Note next that if, in each row of M', we push all the 1's to the left until they are "bumper to bumper" and likewise push all the zeros to the right, then the number of 1's in the first k column of the resulting matrix will be given by the expression

$$c_1(\sigma^{-1}) \wedge k + c_2(\sigma^{-1}) \wedge k + \cdots + c_n(\sigma^{-1}) \wedge k$$

where for convenience we set $a \wedge b = min(a, b)$. This is simply due to the fact that in row i of M' there are $c_i(\sigma^{-1})$ 1's altogether and we can't fit more than k in the first k columns of any row. Consequently we must have

$$c_{j_1}(\sigma) + c_{j_2}(\sigma) + \cdots + c_{j_k}(\sigma) \;\leq\; c_1(\sigma^{-1}) \wedge k + c_2(\sigma^{-1}) \wedge k + \cdots + c_n(\sigma^{-1}) \wedge k \tag{2.3.8}$$

Note further that if $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ is any partition and $\mu' = (\mu'_1, \mu'_2, \ldots, \mu'_{m'})$ is its conjugate we necessarily have for any $k \leq m'$

$$\mu_1 \wedge k + \mu_2 \wedge k + \cdots + \mu_m \wedge k \;=\; \mu'_1 + \mu'_2 + \cdots + \mu'_k$$

Thus if k is less than the number of parts in both $\lambda(\sigma)$ and $\lambda'(\sigma^{-1})$ we may rewrite 2.3.8 as

$$\lambda_1(\sigma) + \lambda_2(\sigma) + \cdots + \lambda_k(\sigma) \;\leq\; \lambda'_1(\sigma^{-1}) + \lambda'_2(\sigma^{-1}) + \cdots + \lambda'_k(\sigma^{-1})\,.$$

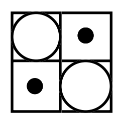
This shows that $\lambda(\sigma)$ is dominated by $\lambda'(\sigma^{-1})$. To prove the last assertion, note that equality in 2.3.8 for all k can only hold true if and only if none of the 1's have moved. Let us take a moment to find out when can this happen. Note first that the way we constructed M', it follows that the 1's in the r^{th} column of M' are in the rows indexed by the elements of $C_{j_r}(\sigma)$ (†). This given, we claim that no motion of 1's forces the set inclusions

$$C_{j_1}(\sigma) \supseteq C_{j_2}(\sigma) \supseteq \cdots \supseteq C_{j_n}(\sigma) \quad . \tag{2.3.9}$$

Indeed if any of these containements did not hold then there would be a 1 to the right of a 0 in M' and that 1 would move. This show that equality in 2.3.7 implies that the components of the code sequence $C(\sigma)$ are totally ordered by inclusion. Conversely, if this holds true, then the permutation $j_1 j_2 \cdots j_n$ that yields $c_{j_1}(\sigma) \geq c_{j_2}(\sigma) \geq \cdots \geq c_{j_n}(\sigma) \geq$ will necessarily produce 2.39 as well, and under these conditions there would be no possible movement of 1's, forcing equality in 2.38 for all k and equality in 2.3.7 as well. This completes our proof.

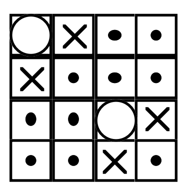
Remark 2.3.2

We should point out that if the sets $C_j(\sigma)$ are not totally ordered by inclusion if and only if there are a pair of indices $r < s$ for which both containements $C_r(\sigma) \subseteq C_s(\sigma)$ and $C_s(\sigma) \subseteq C_r(\sigma)$ simultaneously fail. However this will happen if and only if in the columns r and s the circle diagram $CD(\sigma)$ contains a 2×2 subdiagram of the form



2.3.10

note further that locating the two "X"'s that cause these "•" and the "X"'s to the right and below the second "O" we will necessarily find in $CD(\sigma)$ a 4×4 subdiagram of the form

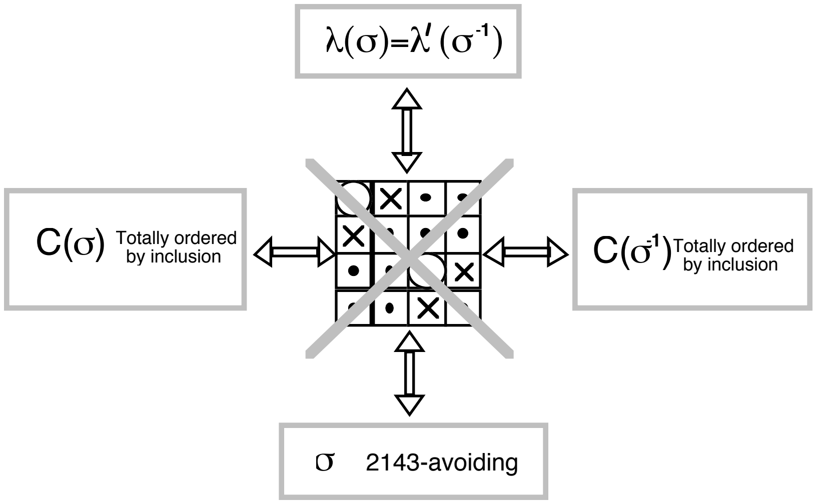


2.3.11

This given, we arrive at the conclusion that equality holds in 2.37 if and only if $CD(\sigma)$ contains no such 4×4 subdiagrams. Finally we should add that what we did with $C(\sigma)$ we could just as well have done with $C(\sigma^{-1})$. Thus we see that the esclusion of 4×4 subdiagrams of the form in 2.3.11 is also equivalent to $C(\sigma^{-1})$ being totally ordered by inclusion. Adding the notion of pattern avoidance to all this we may schematically represent the

(†) see definition 2.3.1

contents of this remark by the following diagram of equivalences.



2.3.12

This brings us to another remarkable class of permutations:

Definition 2.3.2

We say that σ is "Vexillary" if and only if it satisfies any of the equivalent conditions displayed in 2.3.12.

We should note that this terminology is due to Lascoux-Schützenberger who apparently used the prefix *"Vexill"* to express the presence of the "flag" of subsets we see in 2.3.9.

Remark 2.3.3

Note that if a permutation σ has a 2143 subpattern then it has also a 132 subpattern. Moreover between the "2" and the "1" σ will necessary have a have a descent and likewise between the "4" and the "3" it will have another descent. This brings us two important subclasses of Vexillary permutations that play a crucial role in the study of reduced decompositions.

Definition 2.3.3

A 132-avoiding permutation will be called "Dominant "

Definition 2.3.4

A permutation with only one descent will be called "Grassmanian"

These two classes of permutations have further useful characterizations.

Proposition 2.3.2

For a permutation $\sigma \in S_n$ the following conditions are equivalent

(i) σ is dominant

(ii) The circles in $CD(\sigma)$ fill an english Ferrers diagram.

(iii) The code sequence $C(\sigma)$ is decreasing.

(iv) The code of σ is weakly decreasing.

Proof

Note that from the definitions in 2.3.1 and 2.3.3 we derive that (ii) is equivalent to the condition

$$C_i(\sigma) \;=\; \{1,2,\cdots,c_i(\sigma)\} \qquad \text{with} \;\; c_i(\sigma) \geq c_{i+1}(\sigma) \;\; \text{for} \;\; i=1,2,\ldots,n-1 \tag{2.3.13}$$

Thus (ii) implies

$$C_1(\sigma) \supseteq C_2(\sigma) \supseteq C_3(\sigma) \supseteq \cdots \supseteq C_{n-1}(\sigma)\ .$$

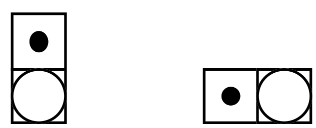
Consequently $(ii)\rightarrow(iii)\rightarrow(iv)$. We next prove $(iv)\rightarrow(ii)$ by showing that the condition

$$c_i(\sigma) \geq c_{i+1}(\sigma) \;\; \text{for} \;\; i=1,2,\ldots,n-1 \tag{2.3.14}$$

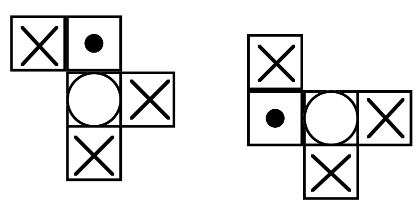
implies 2.3.13. We proceed by induction on "i". Clearly, in any case we have

$$C_1(\sigma) \;=\; \{1,2,\cdots,c_1(\sigma)\}\ .$$

Now note that if $C_i(\sigma)=\{1,2,\cdots,c_i(\sigma)\}$ then all the "X" 's in the first $c_i(\sigma)$ rows of $CD(\sigma)$ must be in columns $i+1,i+2,\cdots,n$. Thus $C_{i+1}(\sigma)=\{1,2,\cdots,k\}$ if the "X" in column $i+1$ is in a row $k \leq c_i(\sigma)$. But then $|C_{i+1}(\sigma)| = c_{i+1}(\sigma) \leq c_i(\sigma)$ forces $k=c_{i+1}(\sigma)$ and completes the induction. Thus (ii), (iii) and (iv) are equivalent. To complete the argument we show that 132-avoiding is equivalent to (ii). To this end note that an english Ferrers diagram λ is characterized by the property that all cells NORTH or WEST of a cell of λ are in λ. Now if one of these conditions fails for $CD(\sigma)$ it necessarily follows that $CD(\sigma)$ must contain one of the subpatterns below



However, if we add the "X" that causes the "$\bullet$" and add the "X" 's that are in the column and row of the "O" we see that $CD(\sigma$ would necessarily contain one of the patterns below



forcing σ to have a 132 subpattern. Conversely, we easily see that the presence of a "132" in σ would prevent the circles of $CD((\sigma)$ to form a Ferrers' diagram. In summary we see that "*not* (i)" is equivalent to "*not* (ii)". This proves that (i),(ii),(iii) and (iv) are equivalent as asserted.

Proposition 2.3.3

For a permutation $\sigma \in S_n$ the following conditions are equivalent

(i) σ is Grassmanian with descent at r

(ii) $c_1(\sigma) \leq c_2(\sigma) \leq \cdots \leq c_r(\sigma) > 0$ and $c_i(\sigma)=0$ for all $i>r$.

Proof

Note that for any permutation σ we have

$$\begin{aligned} &a) \quad c_i(\sigma) > c_{i+1}(\sigma) \quad \Longleftrightarrow \quad \sigma_i > \sigma_{i+1} \\ &b) \quad c_i(\sigma) \leq c_{i+1}(\sigma) \quad \Longleftrightarrow \quad \sigma_i < \sigma_{i+1} \end{aligned} \tag{2.3.15}$$

The reason for this is simple. If $\sigma_i > \sigma_{i+1}$, then all the σ_j less than σ_{i+1} to the right of σ_{i+1} are also less than σ_i. Accounting for σ_{i+1} itself, this gives $c_i(\sigma) \geq 1 + c_{i+1}(\sigma)$. Conversely, if $\sigma_i < \sigma_{i+1}$, then all the σ_j less than σ_i that are to the right of σ_i must also be to the right of σ_{i+1}. Thus, in this case, we must have $c_{i+1}(\sigma) \geq c_i(\sigma)$. This given, we see that the condition

$$\sigma_1 < \sigma_2 < \cdots < \sigma_r > \sigma_{r+1} \quad \text{and} \quad \sigma_{r+1} < \sigma_{r+2} < \cdots < \sigma_n$$

is equivalent to

$$c_1(\sigma) \leq c_2(\sigma) \leq \cdots \leq c_r(\sigma) > 0 \quad \text{and} \quad c_{r+1}(\sigma) = c_{r+2}(\sigma) = \cdots = c_n(\sigma) = 0\ .$$

This proves the proposition.

Remark 2.3.4

Note that if a permutation σ has a 321-subpattern, then it must have at least 2 descents. Indeed, σ will necessarily have descents between the "3" and the "2" and between the "2" and the "1". Thus we see that Grassmanian permutations are also 321-avoiding.

This observation yields us a beautiful corollary of Theorem 2.3.1. It may be stated as follows.

Theorem 2.3.2

If σ is Grassmanian of shape λ then

(i) For every $w \in RED(\sigma)$ the balanced filling T_w of $CD(\sigma)$ can be obtained from a corresponding standard filling τ_w of the Ferrers diagram of λ'.

(ii) Under this correspondence the descent sets $D(w)$ and $D(\tau_w)$ are reversed. That is we have $D(w) = n - D(\tau_w)$.

Proof

From Remark 2.3.4 and (i) of Theorem 2.3.1 it follows that the circles of $C(\sigma)$ fill the cells of a French skew diagram D. However, since σ is also vexillary from (ii) of Proposition 2.3.3 we derive that its code sequence is an increasing sequence of subsets. This forces D to be a "reversed" Ferrers diagram. More precisely, if $\lambda(\sigma) = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ then the columns of D will have lengths $\lambda_k, \lambda_{k-1}, \ldots, \lambda_1$, and its rows will have lengths $\lambda'_1, \lambda'_2, \ldots, \lambda'_h$ with $\lambda' = (\lambda'_1, \lambda'_2, \ldots, \lambda'_h)$ the conjugate of $\lambda(\sigma)$. This means that if we rotate D 180 degrees, we will obtain precisely the Ferrers diagram of the partition λ'. In particular, this rotation gives a correspondence between the standard fillins of D and the standard fillings of the Ferrers diagram of λ'. To do this we only need to replace, after rotation, each label k by its complement $n + 1 - k$. This given, for $w \in RED(\sigma)$, let T_w be the corresponding standard labeling of $CD(\sigma)$, τ_w be the induced standard labeling of D and finally let τ'_w be the standard labeling of the Ferrers diagram of λ' that we obtain by rotating and complementing τ_w. It is easily seen that under the mapping $\tau_w \longrightarrow \tau'_w$ an element i of $D(\tau_w)$ is sent onto the element $n - i$ of $D(\tau'_w)$. Thus Part (ii) of this theorem follows from (ii) and (iii) of Theorem 2.3.1.

From Theorems 2.3.1 and 2.3.4 we may derive two remarkable identities which essentially go back to R. Stanley's original paper. To state them we need to introduce some notation. To begin with, it will be convenient to use compositions to represent descent sets. More precisely, given a subset

$$S = \{1 \leq i_1 < i_2 < \cdots < i_k < n\} \ \subseteq \ [1,n]$$

we set

$$p(S,n) \quad = \quad [i_1 \ , \ i_2 - i_1 \ , \ i_3 - i_2 \ , \ \ldots \ , \ i_k - i_{k-1} \ , \ n - i_k] \qquad 2.3.16$$

Note that from this notation not only we can recover S but also the interval $[1,n]$ we are considering S a subset of. This given, for any word $w = a_1 a_2 \cdots a_l$ we shall here and after set

$$p(w) \quad = \quad p\big(D(w), n\big) \qquad 2.3.17$$

For instance for

$$w = 23453624 \in RED([1,5,3,6,4,7,2])$$

we have

$$D(w) \quad = \quad \{4,6\} \subseteq [1,8]$$

Thus

$$p(w) \quad = \quad [4,2,2] \ .$$

In the same vein for a standard labeling τ of a french or english skew or straight Ferrers diagram on $1,2,\ldots,n$ we set

$$p(\tau) \quad = \quad p\big(D(\tau), n\big) \ . \qquad 2.3.18$$

For instance for the french standard tableau

$$\tau \quad = \quad \begin{array}{ccc} 4 & 8 & \\ \underline{3} & 5 & \underline{7} \\ 1 & \underline{2} & \underline{6} \end{array}$$

the underlined elements are its "descents" thus

$$p(\tau) \quad = \quad [2,1,3,1,1]$$

Using this notation we can represent collection of "descent" sets by formal sums of variables indexed by compositions. More precisely we set for a given $\sigma \in S_n$

$$\Xi(\sigma) \quad = \quad \sum_{w \,\in\, RED(\sigma)} x_{p(w)} \qquad 2.3.19$$

For example we have

$$RED([3,4,2,1) \quad = \quad \big\{[1,2,3,1,2] \ , \ [1,2,1,3,2] \ , \ , [2,1,2,3,2] \ , \ [2,3,1,2,3] \ , \ [2,1,3,2,3]\big\}$$

from which we deduce that

$$\Xi([4,3,1,2]) \quad = \quad x_{32} + x_{221} + x_{131} + x_{23} + x_{122} \qquad 2.3.20$$

In the same vein, for a french or english straight or skew Ferrers diagram D we set

$$\Sigma(D) \;=\; \sum_{\tau\in ST(D)} x_{p(\tau)} \qquad 2.3.21$$

where this summation is over all standard labelings of D.

For instance, the standard tableaux of shape $(3,2)$ with descents underlined are

$$\begin{matrix}2&4&\\ \underline{1}&\underline{3}&5\end{matrix}\;,\quad \begin{matrix}2&5&\\ \underline{1}&3&\underline{4}\end{matrix}\;,\quad \begin{matrix}3&4&\\ 1&\underline{2}&5\end{matrix}\;,\quad \begin{matrix}3&5&\\ 1&\underline{2}&\underline{4}\end{matrix}\;,\quad \begin{matrix}4&5&\\ 1&2&\underline{3}\end{matrix}\;.$$

and this gives

$$\Sigma([3,2]) \;=\; x_{122}+x_{131}+x_{23}+x_{221}+x_{32}\;.$$

The fact that we get the same expression here as in in 2.3.20 is not an accident. Indeed, it is a particular case of the main result proved by Stanley's in []. We can show now that it is a consequence of Theorem 2.3.2.

In fact, Theorems 2.3.1 and 2.3.2 yield us the following two general results.

Theorem 2.3.3

(1) If σ is 321-avoiding with associated french skew diagram D then

$$\Xi(\sigma) \;=\; \Sigma(D)\;, \qquad 2.3.22$$

(2) If σ is Grassmanian of shape λ and we let λ also denote the Ferrers diagram of shape λ then

$$\Xi(\sigma) \;=\; \Sigma(\lambda')\;. \qquad 2.3.23$$

Proof

The identity in 2.3.22 is simply another way of stating part (iii) of Theorem 2.3.1. Now from part (ii) of Theorem 2.3.2 we derive that if σ is Grassmanian then

$$\Xi(\sigma) \;=\; \sum_{\tau\in ST(\lambda)} x_{p^*(\tau)}$$

where, for a composition $p=(p_1,p_2,\ldots,p_r)$ we set p^* denotes the reversed composition $p^*=(p_r,\ldots,p_2,p_1)$. But then 2.3.23 follows from the fact that for any Ferrers diagram we have

$$\sum_{\tau\in ST(\lambda)} x_{p^*(\tau)} \;=\; \sum_{\tau\in ST(\lambda)} x_{p(\tau)}\;.$$

It turns out that Grassmanian permutations are also closely related to dominant permutations. More precisely we have

Proposition 2.3.4

Let $\sigma=\sigma_1\sigma_2\cdots\sigma_n$ be Grassmanian with descent at r, and shape $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_r)$ then

$$\sigma' \;=\; \sigma_r\sigma_{r-1}\cdots\sigma_1\sigma_{r+1}\sigma_{r+2}\cdots\sigma_n \qquad 2.3.24$$

is dominant of shape

$$\mu=(\lambda_1+r-1,\lambda_2+r-2,\ldots,\lambda_r+r-r) \qquad 2.3.25$$

Proof

If

$$\sigma_1 < \sigma_2 < \cdots < \sigma_r > \sigma_{r+1} \quad \text{and} \quad \sigma_{r+1} < \sigma_{r+2} < \cdots < \sigma_n$$

then the code of σ is

$$c(\sigma) = (\sigma_1 - 1, \sigma_2 - 2, \ldots, \sigma_r - r, 0, 0, \ldots, 0)$$

and

$$\lambda(\sigma) = (\sigma_r - r, \ldots, \sigma_2 - 2, \sigma_1 - 1) \; . \tag{2.3.26}$$

On the other hand, we derive from 2.3.24 that

$$\lambda(\sigma') = (\sigma_r - 1, \ldots, \sigma_2 - 1, \sigma_1 - 1) \tag{2.3.27}$$

and (iv) of Proposition 2.3.2 gives that σ' is dominant. The final assertion in 2.3.25 follows by comparing 2.3.26 and 2.3.27.

Now it develops that collections of Grassmanian permutations and in particular also collections vexillary permutations can be used to encode certain characteristics of general permutations. This remarkable discovery of Lascoux and Schützenberger will be the main topic of the next section.

2.4 The Lascoux-Schützenberger tree of a general permutation.

Before we can proceed with the construction of this tree we need to review a few basic facts about the so called "*Bruhat*" partial orders. To begin, let us use the symbol t_{ij} to denote the transposition (i,j). Here and after we shall use the symbol "t" to refer to a generic such transposition and reserve the letter s to refer to a generic simple transpositions $s_i = (i, i+1)$. We also set

$$\mathcal{T} = \mathcal{T}_n = \{t_{ij} \,:\, 1 \le i < j \le n\} \quad \text{and} \quad \mathcal{S} = \mathcal{S}_n = \{\, s_i\,;\, i = 1, 2, \ldots, n-1\,\} \qquad 2.4.1$$

Note that if

$$\sigma' = \sigma \times t \quad \text{with} \quad t \in \mathcal{T} \qquad 2.4.2$$

then

$$\sigma' = t' \times \sigma \quad \text{with} \quad t' \in \mathcal{T} \qquad 2.4.3$$

Indeed from 2.4.2 we derive that

$$t' \;=\; \sigma \times t \times \sigma^{-1}$$

In other words, if 2.4.3 holds with $t = t_{ij}$ then 2.4.3 holds with $t' = t_{\sigma_i, \sigma_j}$. Keeping this observation in mind we set

$$\sigma{-}B{\rightarrow}\sigma' \quad \Longleftrightarrow \quad \begin{cases} a) \quad \sigma' = \sigma \times t \quad \text{with } t \in \mathcal{T} \\ b) \quad l(\sigma') > l(\sigma) \end{cases} \qquad 2.4.4$$

Note that if $t = t_{ij}$ we see that $b)$ simply says that $\sigma_i < \sigma_j$. Note further that when $\sigma_i < \sigma_j$ we have

$$l(\sigma') = l(\sigma) + 1 \qquad \text{if and only if} \qquad \{\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{j-1}\} \cap [\sigma_i, \sigma_j] \;=\; \emptyset \qquad 2.4.5$$

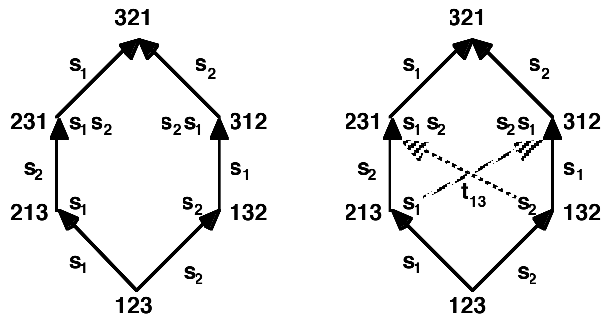
This is simply due to the fact that for any $i < k < j$ such that σ_k is in the interval $[\sigma_i, \sigma_j]$ the number of inversions of σ increases by 2 as we transpose σ_i with σ_j. We shall refer to "$\sigma \;-B\!\rightarrow\; \sigma'$" as a "*Bruhat transition* " and as a "*simple Bruhat transition* " when 2.4.5 holds true. This given, the transitive closure of the relation "$\sigma \;-B\!\rightarrow\; \sigma'$ ", denoted "$<_B$" is usually referred to as "*Bruhat partial order of* S_n".

Remark 2.4.1

We should note that the "*weak Bruhat order*", denoted "$<_W$" is similarly obtained. We call "*weak Bruhat*" transitions interchanges of the form

$$\sigma{-}W{\rightarrow}\sigma' \quad \Longleftrightarrow \quad \begin{cases} a) \quad \sigma' = \sigma \times s \quad \text{with } s \in \mathcal{S} \\ b) \quad l(\sigma') > l(\sigma)\,, \end{cases} \qquad 2.4.6$$

and then define "$<_W$" be the transitive closure of weak Bruhat transitions. With this terminology the reduced decompositions of a permutation $\sigma \in S_n$ may be viewed as the maximal (unrefinable) chains joining the identity of S_n to σ. The following display illustrates the difference betwee the weak and strong Bruhat orders of S_3.



The following is an important tool for working with Bruhat order.

Proposition 2.4.1 (EXCHANGE PROPERTY)

Let σ a permutation of length l and suppose that

$$w = a_1 a_2 \cdots a_l \in RED(\sigma') \tag{2.4.7}$$

let

$$l(\sigma') < l(\sigma) \qquad \text{with} \qquad \sigma' = \sigma t_{rs} \qquad (r < s)\ . \tag{2.4.8}$$

Then for some $i = 1, 2, \ldots, l$ we have

$$a)\ \ \sigma' = s_{a_1} s_{a_2} \cdots s_{a_{i-1}} s_{a_{i+1}} \cdots s_{a_l} \qquad \text{and} \qquad b)\ \ \sigma = s_{a_1} s_{a_2} \cdots s_{a_{i-1}} s_{a_{i+1}} \cdots s_{a_l} t_{rs} \tag{2.4.9}$$

In particular if $l(\sigma') = l(\sigma) - 1$ then we also have

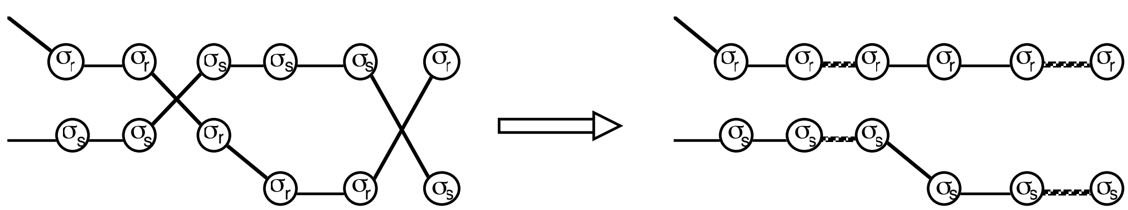
$$w' = a_1 a_2 \cdots a_{i-1} a_{i+1} \cdots a_l \in RED(\sigma') \tag{2.4.10}$$

Proof

The assumption in 2.4.8 says that $\sigma_r > \sigma_s$. This together with 2.47 yields that in the line diagram $\mathcal{M}(a_1 a_2 \cdots a_l)$ the σ_r and σ_s lines cross precisely once. Assuming that this crossing occurs at time i by the action of s_{a_i}, then removing s_{a_i} and t_{rs} from the factorization

$$\sigma' \quad = \quad s_{a_1} s_{a_2} \cdots s_{a_i} \cdots s_{a_l} t_{rs} \tag{2.4.11}$$

we simply obtain the factorization in 2.4.9 a) which will then achieve the same end result. Schematically we may represent the passing from 2.4.11 to 2.4.9 a) as replacing the line diagram on the left by the one on the right in the following display.



Clearly 2.4.9 b) follows from 2.4.9 a) and 2.4.8. Finally, since the factorization in 2.4.9 a) has $l-1$ factors, the assertion in 2.4.10 will necessarily hold true when $l(\sigma')=l-1$. This completes our proof.

Remark 2.4.2

We should note that removing s_{a_i} from a factorization

$$\sigma \;=\; s_{a_1}s_{a_2}\cdots s_{a_i}\cdots s_{a_m}$$

may be simply obtained upon multiplication of σ on the right by the transposition

$$t \;=\; s_{a_m}s_{a_{m-1}}\cdots s_{a_{i+1}}s_{a_i}s_{a_{i+1}}\cdots s_{a_{m-1}}s_{a_m}$$

It will be convenient here and after to denote the omission of a factor by sourrounding it by square brackets. That is we shall write

$$\sigma \;=\; s_{a_1}s_{a_2}\cdots [s_{a_i}]\cdots s_{a_m}$$

for

$$\sigma \;=\; s_{a_1}s_{a_2}\cdots s_{a_{i-1}}s_{a_{i+1}}\cdots s_{a_m}$$

As a corollary of Proposition 2.4.1 we obtain.

Proposition 2.4.2

If the permutation σ has the factorization

$$\sigma \;=\; s_{a_1}s_{a_2}\cdots s_{a_m} \qquad 2.4.12$$

Then indices $1\leq i_1<i_2<\cdots<i_k\leq m$ can be selected so that

$$\sigma \;=\; s_{a_{i_1}}s_{a_{i_2}}\cdots s_{a_{i_k}} \qquad 2.4.13$$

Gives a reduced factorization of σ.

Proof

If $l(\sigma)=m$ then 2.4.12 is reduced and there is nothing to prove. If $l(\sigma)<m$ then as we compute the successive products

$$s_{a_1} \;\rightarrow\; s_{a_1}s_{a_2} \;\rightarrow\; s_{a_1}s_{a_2}s_{a_3} \;\rightarrow\; \cdots$$

sooner or later we will have a drop in length. Letting $j+1$ be the first time this happens we will have

$$l(s_{a_1}s_{a_2}s_{a_3}\cdots s_{a_j})=j \quad \text{and} \quad l(s_{a_1}s_{a_2}s_{a_3}\cdots s_{a_{j+1}})=j-1\ .$$

From the exchange property we then deduce that for some $1\leq i\leq j$ we will have the reduced factorization

$$s_{a_1}s_{a_2}s_{a_3}\cdots s_{a_{j+1}} \;=\; s_{a_1}s_{a_2}\cdots [s_{a_i}]\cdots s_{a_j}$$

Continuing the successive multiplications

$$s_{a_1}s_{a_2}\cdots [s_{a_i}]\cdots s_{a_j} \;\rightarrow\; s_{a_1}s_{a_2}\cdots [s_{a_i}]\cdots s_{a_j}s_{a_{j+2}} \;\rightarrow\; s_{a_1}s_{a_2}\cdots [s_{a_i}]\cdots s_{a_j}s_{a_{j+2}}s_{a_{j+3}} \;\rightarrow\cdots$$

We may run into another drop in length. If this occurs at time $k+1$, another application of the exchange property will yield a reduced factorization of the form

$$s_{a_1}s_{a_2}\cdots s_{a_{k+1}} \;=\; s_{a_1}s_{a_2}\cdots[s_{a_i}]\cdots[s_{a_{j+1}}]\cdots[s_{a_r}]\cdots s_{a_k}$$

We can easily see that if we carry out this process to completion we will end up obtaining a reduced factorization for σ from an appropriate subword of $a_1a_2\cdots a_m$ precisely as asserted.

The construction of the Lascoux-Schützenberger trees, here and after briefly referred to as "LS-trees", depends on performing certain "*down-up*" transitions of the form

$$\sigma \;\longrightarrow\; u \;\longrightarrow\; \sigma' \qquad 2.4.14$$

where for some $i<r<s$ we have

$$a)\quad \begin{cases} u=\sigma\times t_{rs} \\ l(u)=l(\sigma)-1 \end{cases} \quad \text{and} \quad b)\quad \begin{cases} \sigma'=u\times t_{ir} \\ l(\sigma')=l(u)+1 \end{cases} \qquad 2.4.15$$

This given, for a fixed $u\in S_n$ and $1\le r<n$ we set

$$\begin{aligned}\Psi(u,r) &= \big\{\alpha\in S_n:\alpha=u\times t_{rs}\ \&\ l(\alpha)=l(u)+1 \text{ with } s>r\big\},\\ \Phi(u,r) &= \big\{\beta\in S_n:\beta=u\times t_{ir}\ \&\ l(\beta)=l(u)+1 \text{ with } i<r\big\}.\end{aligned} \qquad 2.4.16$$

Now we have the following truly remarkable identity

Theorem 2.4.1

For every $1<r<n$ for which both $\Psi(u,r)$ and $\Phi(u,r)$ are not empty we have

$$\sum_{\alpha\in\Psi(u,r)}\Xi(\alpha) \;=\; \sum_{\beta\in\Phi(u,r)}\Xi(\beta) \qquad 2.4.17$$

The proof of this result will be given in section x.y. In this section we shall start by showing that it naturally leads to LS trees and then derive a number of its important consequences. To this end note that 2.4.17 takes a most interesting form when $\Psi(u,r)$ or $\Phi(u,r)$ contains a single element. The case when $|\Psi(u,r)|=1$ can be stated as follows.

Theorem 2.4.2

Let $\sigma=\sigma_1\sigma_2\cdots\sigma_n\in S_n$ and suppose that for a pair $1<r<s\le n$ the permutation $u=\sigma t_{rs}$ satisfies

(1) $l(u)=l(\sigma)-1$.
(2) $\Psi(u,r) \;=\; \{\sigma\}$
(3) $\Phi(u,r)\neq\emptyset$

Then

$$\Xi(\sigma)=\sum_{\sigma'\in\Phi(u,r)}\Xi(\sigma') \qquad 2.4.18$$

This identity suggests an algorithm for computing the polynomials $\Xi(\sigma)$. The idea is that recursive applications of 2.4.18 should enable us to reduce $\Xi(\sigma)$ to a sum of $\Xi(\sigma')$ which we already know. In view of Theorem 2.3.3, we might hope that we can force all the σ''s occurring in the final sum to be Grassmanian or even only 321-avoiding. It develops that Lascoux and Schützenberger in [] devised precisely such an algorithm for the computation of Littlewood-Richardson coefficients. Curiously, their algorithm (in spite of their claims to the contrary) is hopelessly inefficient as compared with well known methods. Nevertheless, unbeknown to them at that time, and unbeknown to many even at this time, the "*tree*" resulting from their algorithm is precisely what is needed for an efficient way to compute the polynomials $\Xi(\sigma)$ as well as proving some the fundamental properties of the Stanley symmetric functions.

We shall see that conditions (1) and (2) of Theorem 2.4.2 are easily assured. The only thing that is needed is a device for assuring condition (3). This is obtained by means of the following "*shift*" operation introduced by Lascoux and Schützenberger. Using Macdonald's notation this operation may be defined by setting for each $\sigma = \sigma_1\sigma_2\cdots\sigma_n \in S_n$ and and integer $m > 0$:

$$1_m \otimes \sigma \;=\; \begin{bmatrix} 1 & 2 & 3 & \cdots & m & 1+m & 2+m & \cdots & n+m \\ 1 & 2 & 3 & \cdots & m & \sigma_1+m & \sigma_2+m & \cdots & \sigma_n+m \end{bmatrix}. \qquad 2.4.19$$

The relevancy of this operation for our purposes derives from the following simple fact:

Proposition 2.4.3

For all $\sigma \in S_n$ and $m \geq 1$ we have

$$\Xi(1_m \otimes \sigma) \;=\; \Xi(\sigma) \qquad 2.4.20$$

Proof

Note that from 2.4.19 we deduce that

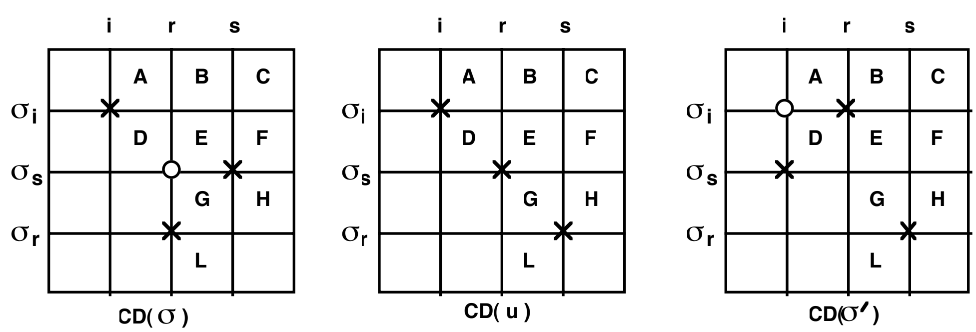
$$a_1a_2\cdots a_l \in RED(\sigma) \iff a_1+m\, a_2+m\,\cdots a_l+m \in RED(\sigma)\,. \qquad 2.4.21$$

Since shifting by a constant each letter of a word does not change its descent set, the identity in 2.4.20 follows then immediately from the definition in 2.3.19.

The following result is basic in assuring that conditions (1) and (2) of Theorem 2.4.2. are satisfied.

Proposition 2.4.4

Let $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ and for a triplet $1 \leq i < r < s \leq n$ suppose that $\sigma_i < \sigma_s < \sigma_r$. Set $u = \sigma \times t_{rs}$ and $\sigma' = u \times t_{ir}$. Decompose the circle diagrams of σ, u and σ' as indicated below



2.4.22

where a letter in a square represents the collection of "X" 's in that open region. Then (a) $\Psi(u,r) = \{\sigma\}$ *and* (b) $\sigma' \in \Phi(u,r)$ *hold true if and only if* $D=\emptyset$, $G=\emptyset$, $H=\emptyset$, $L=\emptyset$.

Proof

To assure that $\sigma \in \Psi(u,r)$ we must have 2.4.15 a). This requires that $\sigma_s < \sigma_r$ and

$$\{\sigma_{r+1}, \sigma_{r+2}, \ldots, \sigma_{s-1}\} \cap [\sigma_r, \sigma_s] = \emptyset \ , \tag{2.4.23}$$

because any element common to these two sets would produce two additional inversions in the transition $u \to \sigma$, violating the second part of 2.4.15 a). Now it is easily seen that 2.4.23 simply means that there are no "X" 's in the open region denoted by G in $CD(\sigma)$, $CD(u)$ and $CD(\sigma')$. Similarly, to assure that $\sigma' \in \Phi(u,r)$ we need to have

$$\{\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{r-1}\} \cap [\sigma_i, \sigma_s] = \emptyset \ . \tag{2.4.24}$$

and this means that there are no "X" 's in the open region denoted by D.

Note further that if we had some $s' > s$ with $\sigma_s < \sigma_{s'} < \sigma_r$ then by taking the one with s' minimal the permutation $u \times t_{rs'}$ would yield us another element of $\Psi(u,r)$. So to satisfy the uniqueness part of condition (a) we must also require that there be no "X" 's in the open regions denoted by H. Likewise if we had an s' with $r < s' < s$ and $\sigma_{s'} > \sigma_r$ then by taking the one with s' smallest the permutation $u \times t_{rs'}$ would yield us another element of $\Psi(u,r)$. This accounts for the requirement $L = \emptyset$ in $CD(\sigma)$, $CD(u)$ and $CD(\sigma')$. This completes our argument.

To complete the picture we need to find out for which permutations $\sigma = \sigma_1, \sigma_2 \cdots \sigma_n$ we can find at least one triplet of indices $1 \leq i < r < s \leq n$ for which the conditions of Proposition 2.4.4 are satisfied. Lascoux and Schützenberger noted the following very simple solution to this problem.

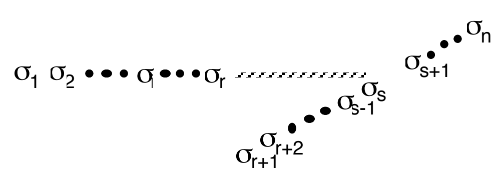
Theorem 2.4.3

If we choose r *to be the last descent of* $\sigma = \sigma_1, \sigma_2 \cdots \sigma_n$ *and let* $s > r$ *be the largest index such that* $\sigma_s < \sigma_r$, *then setting* $u = \sigma \times t_{rs}$ *we shall have* $\Psi(u,r) = \{\sigma\}$ *and there will be at least one index* $i < r$ *for which* $\sigma' = u \times t_{ir} \in \Phi(u,r)$ *provided*

$$\min\{\sigma_j \ : \ j < r\} < \sigma_s \tag{2.4.25}$$

Proof

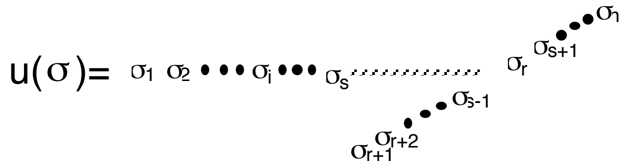
To help visualizing these choices of r and s, in the figure below, we have schematically depicted the behaviour of σ after its last descent.



In other words we have assured the inequalities

$$\sigma_{r+1} < \sigma_{r+2} < \cdots < \sigma_{s-1} < \sigma_s < \sigma_r < \sigma_{s+1} < \sigma_{s+2} < \cdots < \sigma_n \tag{2.4.26}$$

In the same vein the permutation $u = \sigma \times t_{rs}$ which here and after is denoted "$u(\sigma)$", may be depicted as indicated below



2.4.27

Now it is not difficult to see that the inequalities in 2.4.26 guarantee the conditions $G = H = L = \emptyset$ for $CD(\sigma)$ and $CD(u(\sigma))$ assuring that $\Psi(u(\sigma), r) = \{\sigma\}$. Now, if the condition in 2.4.25 is satisfied then by chosing the largest $i < r$ for which $\sigma_i < \sigma_s$ we will have

$$\{\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{r-1}\} \cap [\sigma_i, \sigma_s] \;\;=\;\; \emptyset$$

assuring that $\sigma' = u(\sigma) \times t_{ir} \in \Phi(u(\sigma), r)$. This completes our argument.

This result shows that when condition 2.4.25 is satisfied we are able to express $\Xi(\sigma)$ as in 2.4.18 with $u = u(\sigma)$. But what are we to do if

$$\sigma_1, \sigma_2, \ldots, \sigma_{r-1} \;>\; \sigma_s \tag{2.4.28}$$

Lascoux and Schützenberger have a simple answer also in this case: They simply replace σ by

$$1 \otimes \sigma \;\;=\;\; \begin{bmatrix} 1 & 2 & 3 & \cdots & n+1 \\ 1 & 1+\sigma_1 & 1+\sigma_2 & \cdots & 1+\sigma_n \end{bmatrix}$$

Indeed, since the last descent of $1 \otimes \sigma$ is now at $r+1$ and

$$u(1 \otimes \sigma) \;\;=\;\; 1 \otimes u(\sigma)$$

we can easily see that we have

$$l(u(1 \otimes \sigma)) = l(1 \otimes \sigma) - 1$$

as well as

$$\Psi\big(u(1 \otimes \sigma), r+1\big) \;\;=\;\; \big\{1 \otimes \sigma\big\}$$

Now the inequalities in 2.4.28 can also be rewritten as

$$1 + \sigma_1, 1 + \sigma_2, \ldots, 1 + \sigma_{r-1} \;\;>\;\; 1 + \sigma_s$$

and these yield that the permutation

$$\sigma' \;\;=\;\; u(1 \otimes \sigma) \times t_{1,r+1} \tag{2.4.29}$$

belongs to the set

$$\Phi(u(1 \otimes \sigma), r+1)\,.$$

Moreover, it is easy to see that σ' is the only element of this set. Thus we can apply Theorem 2.4.2 to this case and derive from 2.4.18 that

$$\Xi(\sigma) \;\;=\;\; \Xi(u(1 \otimes \sigma) \times t_{1,r+1}) \tag{2.4.30}$$

We now have all the ingredients we need for the construction of the LS trees.

The Branching Process for $\sigma = \sigma_1\sigma_2\cdots\sigma_n$**:**

Step 1: Locate the last descent of σ*. If this occurs at* r *then let* s *be the largest index such that* $s > r$ *and* $\sigma_s < \sigma_r$ *.*

Step 2: Let $u = \sigma \times t_{rs} = \sigma_1, \sigma_2 \ldots \sigma_{r-1}\,\sigma_s\,\sigma_{r+1}\cdots\sigma_{s-1}\,\sigma_r\,\sigma_{s+1}\cdots\sigma_n$.

Step 3: Case a) If $\Phi(u,r) \neq \emptyset$ *then the children of* σ *are the permutations* $\sigma' \in \Phi(u,r)$ *.*
Case b) If $\Phi(u,r) = \emptyset$ *then* σ *has only one child, namely* $\sigma' = u(1\otimes\sigma)\times t_{1.r+1}$.

Definition 2.4.1

The LS tree of a permutation σ *is the tree obtained by recursive calls of the branching process described above starting with* σ *and stopping the recursion at every child that is Grassmanian.*

To show that this construction always yields a finite tree, Lascoux and Schützenbereger produce the following beautiful estimate for the length of any downward path in the LS tree of a permutation.

Proposition 2.4.5

Let σ *be a permutation of length* l *and let* $d_o(\sigma)$ *and* $d_1(\sigma)$ *denote the first and last descents of* σ*. Assume that for the following chain of permutations*

$$\sigma = \sigma^{(1)} \rightarrow \sigma^{(2)} \rightarrow \cdots \rightarrow \sigma^{(N)}$$

we have

(a) Each is a child of the previous one,
(b) None of them is Grassmanian,

then

$$N \leq l \times \big(d_1(\sigma) - d_o(\sigma)\big) \qquad 2.4.33$$

Before we prove this result it will be good to experiment with the construction of a number of LS trees and understand how simple the process really is.

To begin let us make more explicit our construction of the children of σ. To this end note that in Case a) the children of σ are the permutations $\sigma' = u \times t_{ir}$ for each $i < r$ such that

$$\sigma_i < \sigma_s \quad \& \quad \big\{\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{r-1}\big\} \cap [\,\sigma_i\,,\ \sigma_s\,] \;=\; \emptyset\,,$$

and when this holds

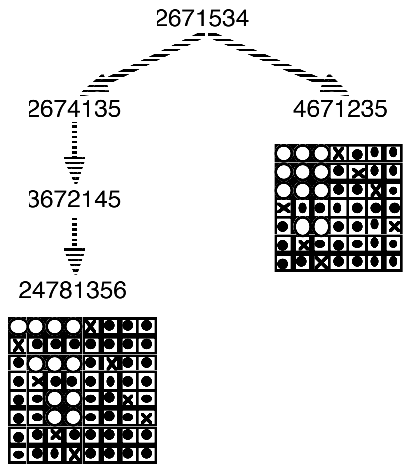
$$\sigma' \;=\; \sigma_1, \sigma_2 \cdots \sigma_{i-1}\,\sigma_s\,\sigma_{i+1}\cdots\ldots\sigma_{r-1}\,\sigma_i\,\sigma_{r+1}\cdots\sigma_{s-1}\,\sigma_r\,\sigma_{s+1}\cdots\sigma_n. \qquad 2.4.34$$

In Case b), the unique child $\sigma' = u(1\otimes\sigma)\times t_{1,r+1}$ is none other than the permutation

$$\sigma' = \begin{bmatrix} 1 & 2 & 3 & \cdots & r & r+1 & r+2 & \cdots & s & s+1 & s+2 & \cdots n+1 \\ \underline{\sigma_s} & \underline{\sigma_1} & \underline{\sigma_2} & \cdots & \underline{\sigma_{r-1}} & 1 & \underline{\sigma_{r+1}} & \cdots & \underline{\sigma_{s-1}} & \underline{\sigma_r} & \underline{\sigma_{s+1}} & \cdots \underline{\sigma_n} \end{bmatrix} \qquad 2.4.35$$

with the understanding that $\underline{x} = x + 1$.

For our first example we take the permutation $\sigma = 2671536$, which has a simple but not entirely trivial LS tree. In the figure below we depict this tree with a circle diagrams appended at each leaf.



Note then that multiple applications of the identities in 2.4.18 and 2.4.30 give us the relations

$$\Xi(2671534) = \Xi(2674135) + \Xi(4671235)$$
$$\Xi(2674135) = \Xi(3672145)$$
$$\Xi(3672145) = \Xi(24781356)$$

On the other hand since 24781346 and 4671235 are Grassmanian of shapes $[4421]$ and $[443]$, from Theorem 2.3.3 we derive that

$$\Xi(24781356) \;=\; \Sigma([4322]) \qquad , \qquad \Xi(4671235) = \Sigma([3332])$$

Combining all these identities we derive that

$$\Xi(2671534) \;=\; \Sigma([4322]) \;+\; \Sigma([3332])$$

Thus, in particular it follows that the number reduced decompositions of 2671534 is equal to the number of standard tableaux of shapes $[4421]$ and $[443]$.

We can easily see from this example that the relations in 2.4.18 and 2.4.30 combined with Theorem 2.3.3 yield us the following general result

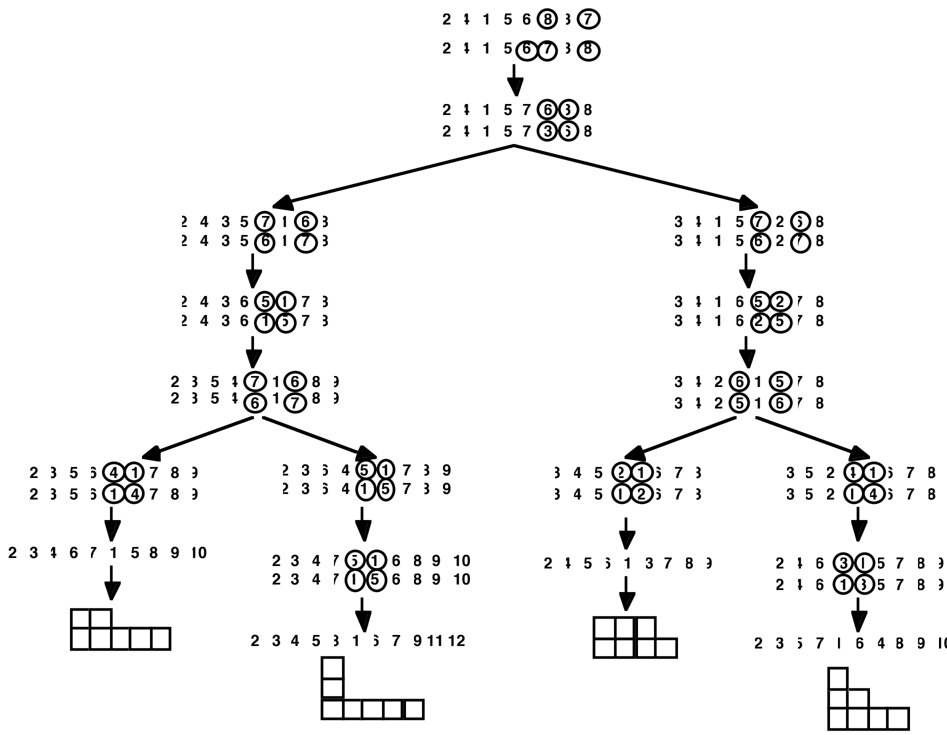
Theorem 2.4.4

On the validity of Theorem 2.4.1 and Proposition 2.4.5, for any permutation σ, we have the expansion

$$\Xi(\sigma) \;=\; \sum_{\sigma' \in LeavesLS(\sigma)} \Sigma\big(\lambda'(\sigma')\big) \qquad 2.4.36$$

where the symbol "$\sigma' \in LeavesLS(\sigma)$" is to indicatee that the summation is over the leaves of the LS tree of σ

We shall take as our next example the permutation $\sigma = 24156837$. To follow step by step the operations that yield the LS tree of this permutation, in the display below we have placed under each σ the permutation $u(\sigma)$ or $u(1 \otimes \sigma)$ as the case may be, and right below we display the offspring. The circled indices are those that get transposed as we pass from a σ to its corresponding $u(\sigma)$. Finally under each Grassmanian leaf σ' we draw the Ferrers diagram of the partition $\lambda'(\sigma')$.



Thus from Theorem 2.4.4 we derive that

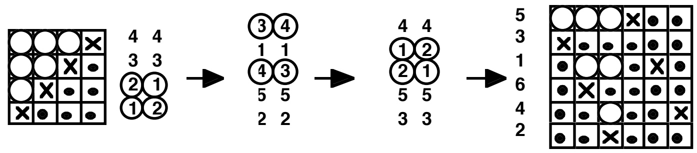
$$\Xi(34156837) = \Sigma([5,2]) + \Sigma([5,1,1]) + \Sigma([4,3]) + \Sigma([4,2,1]) \qquad 2.4.37$$

This example is particularly interesting since the permutation 34156837 is 321-avoiding with corresponding diagram the French skew partition $[5,5,2]/[4,1]$. Now it develops that the skew Schur function $S_{[5,5,2]/[4,1]}$ has the Schur function expansion

$$S_{[5,5,2]/[4,1]} = S_{[5,2]} + S_{[5,1,1]} + S_{[4,3]} + S_{[4,2,1]} \qquad 2.4.38$$

The fact that the right hand sides of 2.4.37 and 2.4.38 are essentially identical is not an accident. In fact it is only an instance of the general fact discovered by Lascoux and Schützenberger that the LS tree can be used to compute Littlewood-Richardson coefficients.

For the final example we have chosen $\sigma = 4321$. Here, to save space, we have depicted the LS tree horizontally. In particular the permutations must be read from top to bottom. We have depicted the circle diagrams of the starting and ending permutations.



We should notice two important facts. First we see here a case when each parent has a single child. This not an accident. As we shall soon see this is always true for vexillary permutations. Secondly we might guess that for the general reversing permutation

$$\sigma^{(n)} \;=\; \begin{bmatrix} 1 & 2 & \cdots & n-1 & n \\ n & n-1 & \cdots & 2 & 1 \end{bmatrix},$$

which is dominant of shape $[n-1, n-2, \ldots, 2, 1]$, the Grassmanian permutation which is the single leaf of its tree has always an associated French skew diagram obtained as 180^o rotation of of the diagram of $[n-1, n-2, \ldots, 2, 1]$. We leave the proof of this to the reader and derive from Theorem 2.3.3 the following result which essentially goes back to Richard Stanley.

Theorem 2.4.5

For the top permutation $\sigma^{(n)} \in S_n$ we have

$$\Xi(\sigma^{(n)}) \;=\; \Sigma([n-1, n-2, \ldots, 2, 1])$$

In particular the number of reduced decomposition of $\sigma^{(n)}$ is equal to the number of standard tableaux of "staircase" shape $[n-1, n-2, \ldots, 2, 1]$.

Our next task is the proof of Proposition 2.4.5. However before we do this we need some preliminary observations and an auxiliary result. To begin, given a permutation σ, it will be good to distinguish children σ' resulting from Case a) of the branching process from those resulting from Case b). We shall call the former "*regular*" children and the latter "*lateral*" children.

It will be good to order regular children $\sigma' = u(\sigma) \times t_{ir}$ according to increasing i. More generally, under the hypotheses of Proposition 2.4.4 let

$$\Phi(u, r) \;=\; \left\{ u \times t_{i_1,r} \;,\quad u \times t_{i_2,r} \;,\; \cdots \;,\quad u \times t_{i_m,r} \right\}$$

with $i_1 < i_2 < \cdots < i_m$. Then it is easy to see that we must also have

$$\sigma_{i_1} > \sigma_{i_2} > \cdots > \sigma_{i_m} \qquad 2.4.39$$

for otherwise the condition in 2.4.24

$$\{\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{r-1}\} \cap [\sigma_i, \sigma_s] \;=\; \emptyset$$

assuring that $\sigma' = u \times t_{ir} \in \Phi(u, r)$ would be violated. This given we see that i_m must be the last index $i < r$ such that $\sigma_i < \sigma_s$. Following Lascoux and Schützenberger we shall call $u \times t_{i_m,r}$ the "*leader*" of $\Phi(u, r)$.

Remark 2.4.3

For our later purposes it will be good to note that if $\sigma' = u \times t_{i,r}$ is the leader of $\Phi(u,r)$ if and only if

$$\sigma_{i+1}, \sigma_{i+1}, \ldots, \sigma_{r-1} \;\; > \;\; \sigma_s \tag{2.4.40}$$

The following two auxiliary results will provide us with the necessary ingredients for the proof of Proposition 2.4.5.

Lemma 2.4.1

Let $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ and suppose that for a triplet $1 \leq i < r < s \leq n$ we have $\sigma_i < \sigma_s < \sigma_r$. Suppose that for $u = \sigma \times t_{rs}$ and $\sigma' = u \times t_{ir}$ we have

$$\Psi(u,r) = \{\sigma\} \quad , \qquad \sigma' \in \Phi(u,r)\,. \tag{2.4.41}$$

Then

$$\Psi(u^{-1},\sigma_s) = \{\sigma^{-1}\} \quad , \qquad \sigma'^{-1} \in \Phi(u^{-1},\sigma_s) \tag{2.4.42}$$

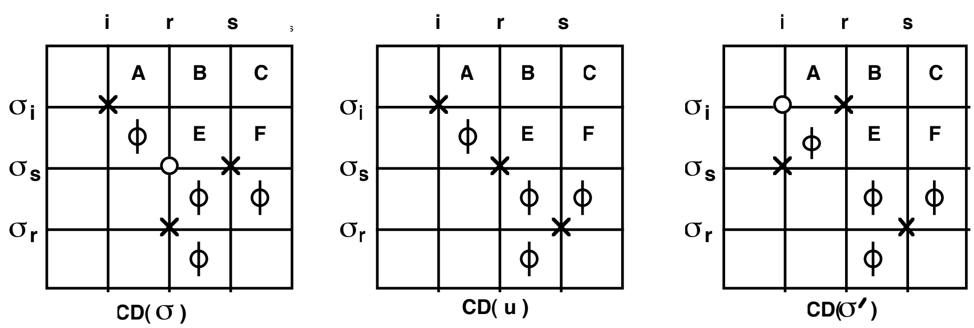
and

$$a) \quad \lambda(\sigma) \leq \lambda(\sigma') \qquad\qquad b) \quad \lambda'(\sigma'^{-1}) \leq \lambda'(\sigma^{-1}) \tag{2.4.43}$$

with equality in a) if and only if σ' is the leader of $\Phi(u,r)$ and equality in b) if and only if σ'^{-1} is the leader of $\Phi(u^{-1},\sigma_s)$.

Proof

From Proposition 2.4.4 we derive that the conditions in 2.4.41 hold if and only if the circle diagrams of σ, u and σ' are of the form given below.



2.4.44

Since the empty sets are symmetrically located with respect to the main diagonal of these diagrams we derive from Proposition 2.4.4 that the conditions in 2.4.41 and 2.4.42 are equivalent. Thus we only need to prove the inequalities in 2.4.43. Letting a,b,c,e,f denote the cardinalty of the sets A,B,C,E,F we immediately deduce, by counting the number of circles in columns i, r and s of $CD(\sigma)$ that

$$c_i(\sigma) = a+b+c \quad , \qquad c_r(\sigma) = b+c+e+f+1 \quad , \qquad c_s(\sigma) = c+f\,. \tag{2.4.45}$$

Doing the same for $CD(\sigma')$ we obtain

$$c_i(\sigma') = a+b+c+1+e+f \quad , \qquad c_r(\sigma') = b+c \quad , \qquad c_s(\sigma') = c+f\,. \tag{2.4.46}$$

Since the changes needed to get $CD(\sigma')$ from $CD(\sigma)$ involve only columns i, r and s we see that we have

$$c_j(\sigma) = c_j(\sigma') \qquad \text{for } j \neq r, s$$

Now we need to distinguish two cases according as

$$a) \quad a > e + f + 1 \qquad \text{or} \qquad b) \quad a \leq e + f + 1\,.$$

In the first case $c_i(\sigma) > c_r(\sigma)$ and 2.4.46 shows that to obtain the Ferrers diagram of $\lambda(\sigma')$ from the Ferrers diagram of $\lambda(\sigma)$ we simply transfer $1 + e + f$ cells from a smaller row to a larger row. In the second case $c_i(\sigma) \leq c_r(\sigma)$ and to obtain the transition $\lambda(\sigma) \rightarrow \lambda(\sigma')$ we need to transfer a cells again from a smaller row to a larger row. So in either case the transfer will cause $\lambda(\sigma')$ to be larger than $\lambda(\sigma)$ in the dominance order. This proves a) of 2.4.43. Now because of 2.4.2, we can apply this very same inequality to the triplet $\sigma^{-1}, u^{-1}, \sigma'^{-1}$ and obtain

$$\lambda(\sigma^{-1}) \leq \lambda(\sigma'^{-1})\,. \tag{2.4.47}$$

This proves 2.4.43 b), since passing to conjugates reverses dominance.

Finally, from 2.4.45 and 2.4.46 we see that in any case we have

$$c_r(\sigma) > c_r(\sigma')\,. \tag{2.4.48}$$

Thus equality in 2.4.43 a) can only occur if and only if $c_r(\sigma) = c_i(\sigma')$ and $c_i(\sigma) = c_r(\sigma')$. This shows that equality holds if and only if $a = 0$. Now a look at the diagrams in 2.4.44 reveals that $a = 0$ occurs if and only if

$$\sigma_j > \sigma_s \quad \forall \quad i < j < r$$

But this is 2.4.40 which, from Remark 2.4.3, is precisely the condition that characterizes a leader of a collection $\Phi(u, r)$. This given, note that since equality in in 2.4.43 b) holds if and only if we have equality in 2.4.47, we see that the finall assertion simply follows by applyng what we have just shown to the triplet $\sigma^{-1}, u^{-1}, \sigma'^{-1}$. This completes the proof.

Lemma 2.4.2

If σ' is a child of a non-Grassmanian σ then

$$d_1(\sigma') - d_o(\sigma') \leq d_1(\sigma) - d_o(\sigma)\,, \tag{2.4.49}$$

and, in case of equality we then have

$$c_{d_1(\sigma')}(\sigma') < c_{d_1(\sigma)}(\sigma)\,. \tag{2.4.50}$$

Proof

Since a lateral child of a permutation σ is a regular child of $1 \otimes \sigma$, and we trivially have

$$\begin{aligned} &a) \quad d_1(1 \otimes \sigma) = d_1(\sigma) + 1 \\ &b) \quad d_o(1 \otimes \sigma) = d_o(\sigma) + 1 \\ &c) \quad c_{d_1(1\otimes\sigma)}(1 \otimes \sigma) = c_{d_1(\sigma)}(\sigma) \end{aligned}$$

we can easily see that we need only prove 2.4,49 and 2.4.50 for regular children. This given, let us assume that σ' is a regular child of σ. Now under this hypothesis we will actually show that

$$a) \quad d_o(\sigma) \leq d_o(\sigma') \quad \text{and} \quad b) \quad d_1(\sigma') \leq d_1(\sigma)\,. \qquad 2.4.51$$

To this end, recall that in this case we have $u = u(\sigma) = \sigma \times t_{rs}$, $\sigma' = u(\sigma) \times t_{ir}$ with $r = d_1(\sigma)$ (the last descent) and again $1 \leq i < r < s \leq n$ with $\sigma_i < \sigma_s < \sigma_r$. In summary

$$\sigma \;=\; \sigma_1 \cdots \sigma_{i-1}\sigma_i\sigma_{i+1} \cdots \sigma_{r-1}\sigma_r\sigma_{r+1} \cdots \sigma_{s-1}\sigma_s\sigma_{s+1} \cdots \sigma_n$$

and

$$\sigma' \;=\; \sigma_1 \cdots \sigma_{i-1}\sigma_s\sigma_{i+1} \cdots \sigma_{r-1}\sigma_i\sigma_{r+1} \cdots \sigma_{s-1}\sigma_r\sigma_{s+1} \cdots \sigma_n$$

In particular

$$\sigma_j = \sigma'_j \qquad \text{for} \quad j \neq i, r, s \;.$$

Thus if $d_o(\sigma) < i-1$ then $d_o(\sigma') = d_o(\sigma)$ and similarly we will have $d_o(\sigma) = d_o(\sigma')$ if $i < d_o(\sigma) < r-1$. If $d_o(\sigma) = i-1$ we may have destroyed the descent at $i-1$ by placing $\sigma_s > \sigma_i$ in position i, giving $d_o(\sigma') > d_o(\sigma)$. Otherwise we again have $d_o(\sigma') = d_o(\sigma)$. If $d_o(\sigma) = i$ then the inequalities

$$\sigma'_i = \sigma_s > \sigma_i > \sigma_{i+1} = \sigma'_{i+1}$$

give $d_o(\sigma') = i$ as well. But what if $d_o(\sigma) > i$ (that is $\sigma_i < \sigma_{i+1}$) and $\sigma_{i+1} < \sigma_s$. Now this cannot happen for otherwise the condition

$$\{\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{r-1}\} \cap [\sigma_i, \sigma_s] \;=\; \emptyset$$

assuring that $\sigma' = u \times t_{ir} \in \Phi(u, r)$ would be violated. Since by assumption σ is not Grassmanian we must have $d_o(\sigma) < r$, thus we are only left to check what happens when $d_o(\sigma) = r-1$. That is if $\sigma_{r-1} > \sigma_r$. However in this case the inequalities $\sigma_r > \sigma_s > \sigma_i$ guarantee that $r-1$ remains a descent as we pass from σ to σ' completing the proof of 2.4.51 a). To prove 2.4.51 b) note that the picture in 2.4.27 clearly shows that neither $u(\sigma)$ nor σ' have a descent after position r. So we only need to check what happens at r itself. To this end note that since $\sigma'_r = \sigma_i$ and $\sigma'_{r+1} = \sigma_{r+1}$ we see that we have $d_1(\sigma') = d_1(\sigma)$ only if the picture is as in 2.4.27 and $\sigma_i > \sigma_{r+1}$. For if $\sigma_i < \sigma_{r+1}$ or worse yet if σ has no elements between positions r and s, (that is if $s = r+1$) then $\sigma'_{r+1} = \sigma_r > \sigma_s > \sigma_i > \sigma'_r$ destroys the descent at r and we will have $d_1(\sigma') < d_1(\sigma)$.

Finally since the computations in the proof of Lemma 2.4.1 apply to the present case as well we see that the inequality in 2.4.48 holds true here with $r = d_1(\sigma)$. In other words we have in any case

$$c_{d_1(\sigma)}(\sigma') < c_{d_1(\sigma)}(\sigma)\,.$$

However, this inequality reduces to 2.4.50 when $d_1(\sigma) = d_1(\sigma')$ and this certainly happens when 2.4.49 reduces to an equality. In fact from 2.4.51 a) and b) we can see that 2.4.49 can be an equality only if we have both $d_1(\sigma') = d_1(\sigma)$ and $d_o(\sigma') = d_o(\sigma)$. This completes our proof.

We now have all the ingredients we need to carry out the final step in the definition of the LS tree.

Proof of Proposition 2.4.5

Let us associate to the member $\sigma^{(i)}$ of the chain

$$\sigma = \sigma^{(1)} \rightarrow \sigma^{(2)} \rightarrow \cdots \rightarrow \sigma^{(N)}$$

the point

$$P^{(i)} \;=\; \big(\, c_{d_1(\sigma^{(i)})}(\sigma^{(i)})\;,\; d_1(\sigma^{(i)}) - d_o(\sigma^{(i)})\,\big) \;\;.$$

Note that, since the components of the code of a permutation never exceed its length, and all the descendants of a permutation have the same length we see that we must have

$$1 \leq c_{d_1(\sigma^{(i)})}(\sigma^{(i)}) \leq l \qquad \text{for } i = 1, 2, \ldots, N\,. \tag{2.4.52}$$

Note further that, since none of the $\sigma^{(i)}$ are Grassmanian, we necessarily have

$$1 \leq d_1(\sigma^{(i)}) - d_o(\sigma^{(i)}) \qquad \text{for } i = 1, 2, \ldots, N\,. \tag{2.4.53}$$

Moreover, we can apply Lemma 2.4.2 to each transition $\sigma^{(i)} \rightarrow \sigma^{(i+1)}$ and, by successive applications of the inequality in 2.49, derive that

$$d_1(\sigma^{(i)}) - d_o(\sigma^{(i)}) \;\leq\; d_1(\sigma) - d_o(\sigma) \qquad \text{for } i = 1, 2, \ldots, N\,. \tag{2.4.54}$$

Combining 2.4.52, 2.4.53 and 2.4.54 we obtain that each of the points $P^{(i)}$ lies in the rectangle

$$S(\sigma) = \big\{ (x, y) \,:\, 0 \leq x \leq l \quad \& \quad 1 \leq y \leq d_1(\sigma) - d_o(\sigma) \,\big\}\,.$$

Since $S(\sigma)$ contains $l \times (d_1(\sigma) - d_o(\sigma))$ lattice points, we see that to prove the inequality in 2.4.33 we need only show that the points $P^{(i)}$ are all distinct. Actually we can do more than that. Indeed, note that from Lemma 2.4.2 we derive that either

$$d_1(\sigma^{(i+1)}) - d_o(\sigma^{(i+1)}) < d_1(\sigma^{(i)}) - d_o(\sigma^{(i)})$$

or

$$d_1(\sigma^{(i+1)}) - d_o(\sigma^{(i+1)}) = d_1(\sigma^{(i)}) - d_o(\sigma^{(i)})$$

but then 2.4.50 gives

$$c_{d_1(\sigma^{(i+1)})}(\sigma^{(i+1)}) < c_{d_1(\sigma^{(i)})}(\sigma^{(i)}) \;\;.$$

Thus the point $P^{(i)}$ keeps moving to the left as it remains in any given row of $S(\sigma)$, This means that after at most l steps we will necessarily have the inequality

$$d_1(\sigma^{(i+1)}) - d_o(\sigma^{(i+1)}) \;<\; d_1(\sigma^{(i)}) - d_o(\sigma^{(i)})$$

which will cause $P^{(i)}$ to skip to a lower row. In summary, we see that $P^{(i)}$, as $i = 1, 2, \ldots, N$, skips from lattice point to lattice point precisely in a strictly decreasing lexicographic manner and thus N cannot exceed the number of lattice points in $S(\sigma)$.

The last result of this section is the following (anticipated) beautiful consequence of Theorem 2.4.4.

Theorem 2.4.6

If σ is vexillary then its LS tree reduces to a chain of vexillary permutations ending with a Grassmanian. In particular it follows that

$$\Xi(\sigma) \;=\; \Sigma\big(\lambda'(\sigma)\big) \tag{2.4.55}$$

Proof

Let σ be any vexillary permutation and let its children be given by the collection

$$\Phi(u,r) \;=\; \big\{u\times t_{i_1,r}\ ,\quad u\times t_{i_2,r}\ ,\ \cdots\ ,\quad u\times t_{i_m,r}\big\}$$

with $r = d_1(\sigma)$ and $i_1 < i_2 < \cdots < i_m$. Then we have seen (2.4.39) that we must also have

$$\sigma_{i_1} > \sigma_{i_2} > \cdots > \sigma_{i_m}$$

However our construction also requires that

$$\sigma_r > \sigma_s > \sigma_{i_1}$$

and the elements $\sigma_{i_1}, \sigma_{i_2}, \sigma_r, \sigma_s$ occur in σ precisely in this order. This means that if $k \geq 2$ then σ would contain a 2143 subpattern which is contrary to our assumption that σ is vexillary. Thus vexillary permutations have only one child, regular or lateral.

Now, recalling (see 2.3.12) that a vexillary permutation σ is characterized by the equality $\lambda(\sigma) = \lambda'(\sigma^{-1})$, we derive from the inequalities in 1.4.43 that the child of a vexillary must also be vexillary and of the same shape as well. This means that the Grassmanian leaf σ' of the LS tree of a vexillary σ will necessarily also have shape $\lambda(\sigma)$. Thus the equality in 2.4.55 is simply another consequence of Theorem 2.3.3.

3. Symmetric Functions and Schubert Polynomials.

3.1 Stanley's Theory of P-Partitions

In these note a partially ordered set (briefly a poset) is a pair $\{\Omega\ ,\ \preceq)$ consisting of a finite set Ω and a partial order "$\preceq$" of the elements of Ω. It will be convenient here and after to let n be the number of elements of Ω. For a given poset $\mathcal{P} = \{\Omega\ ,\ \preceq)$ we let $\mathcal{F}_{\mathcal{P}}$ denote the family of integer valued weakly increasing function of $\mathcal{P}$. In symbols

$$\mathcal{F}_{\mathcal{P}} \;=\; \big\{f\,:\,\Omega\to\mathbb{N}\,:\,x\preceq y\;\Rightarrow\;f(x)\leq f(y)\,\big\}\,.$$

The elements of $\mathcal{F}_{\mathcal{P}}$ are usually referred to as "*$\mathcal{P}$-Partitions*". More generally, given an integral injective labelling ω of Ω we let $\mathcal{F}_{\mathcal{P},\omega}$ denote the subfamily consisting of those elements of $\mathcal{F}_{\mathcal{P}}$ which strictly increase when ω decreases. In symbols

$$\mathcal{F}_{\mathcal{P},\omega} \;=\; \big\{f\in\mathcal{F}_{\mathcal{P}} : x\prec y\ \ \&\ \ \omega_x>\omega_y\;\Rightarrow\;f(x)<f(y)\,\big\}\,. \tag{3.1.1}$$

The elements of $\mathcal{F}_{\mathcal{P},\omega}$ are called "ω-*Compatible* $\mathcal{P}$-*Partitions*".

It will be convenient sometimes to keep these families finite and restrict their elements to take only the values $0, 1, 2, \ldots, N$, for some unspecified very large integer N. In this vein set

$$\mathcal{F}_{\mathcal{P}}(N) \;=\; \big\{ f \in \mathcal{F}_{\mathcal{P}} \,:\, 0 \le f \le N \,\big\} \quad , \quad \mathcal{F}_{\mathcal{P},\omega}(N) \;=\; \big\{ f \in \mathcal{F}_{\mathcal{P},\omega} \,:\, 0 \le f \le N \,\big\} . \qquad 3.1.2$$

This given , to each element $f \in \mathcal{F}_{\mathcal{P},\omega}$ we associate a monomial $x(f)$ in the variables $x_1, x_2, x_3 \ldots$ which is to carry information as to the multiset of values taken by f. More precisely we set

$$x(f) \;=\; \prod_{r \in \Omega} x_{f(r)} \;=\; \prod_{i} x_i^{m_i(f)} \qquad 3.1.3$$

where for $i \in \mathbb{N}$, the integer $m_i(f)$ denotes the number of times f takes the value i.

Extending an idea of MacMahon, Stanley obtained a number of identities concerning the generating functions

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots x_N) \;=\; \sum_{f \in \mathcal{F}_{\mathcal{P},\omega}(N)} x(f) \,. \qquad 3.1.4$$

The main goal of this section is the derivation of some of the identities that are pertinent to our study of reduced decompositions.

The first step is to obtain an expression for $F_{\mathcal{P},\omega}$ that more closely reflects its dependence on the poset $\mathcal{P}$ and its labeling ω. The basic idea is to obtain a decomposition of each element f into a pair $(\sigma(f), p(f))$ consisting of a permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ and a composition $p = (p_1, p_2, \cdots, p_n)$. To this end, let $f \in \mathcal{F}_{\mathcal{P},\omega}(N)$ take the values

$$v_1 < v_2 < \cdots < v_k$$

and set

$$A_i \;=\; \{ x \,:\, f(x) = v_i \,\} \,. \qquad 3.1.5$$

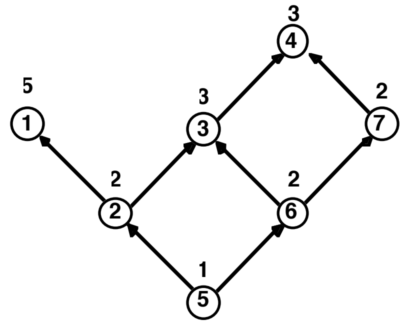
Since Ω has a n elements, there is no loss to assume that the given labeling ω takes the values $1, 2, \ldots, n$. For simplicity it will also be convenient to denote the elements of Ω by their labels. This given, the permutation $\sigma(f)$ is simply obtained by reading the elements of $A_1, A_2, \ldots, A_k$ successively. More precisely we set

$$\sigma(f) \;=\; \uparrow_\omega A_1 \;\uparrow_\omega A_2 \;\cdots\; \uparrow_\omega A_k \,. \qquad 3.1.6$$

Where the symbol "$\uparrow_\omega A_i$" denotes the word obtained by reading the elements of A_i in increasing order. Now, given that $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ we simply set $p(f) = (p_1, p_2, \ldots, p_n)$ with

$$p_1 = f(\sigma_1) \quad \text{and} \quad p_r = f(\sigma_r) - f(\sigma_{r-1}) \quad (\text{for } r = 2, \ldots, n) \,. \qquad 3.1.7$$

This construction is best understood by an example. Let $\mathcal{P}$ be the poset depicted below with the partial order indicated by the arrows and the labeling ω indicated by the integers placed in the circles. We have also given an instance of a particular element $f \in \mathcal{F}_{\mathcal{P},\omega}(N)$ by placing its value above each of the circles.



3.1.8

In this case our definition gives

$$A_1 = \{5\}\ ,\ A_2 = \{2,6,7\}\ ,\ A_3 = \{3,4\}\ ,\ A_4 = \{1\}\,. \qquad 3.1.9$$

thus 3.1.6 gives

$$\sigma(f) \;=\; 5.267.34.1 \qquad 3.1.10$$

Here we have indicated by dots the positions of the descents of the resulting permutation. Following 3.1.7 we then obtain

$$p(f) \;=\; (1, 2-1, 2-2, 2-2, 3-2, 4-3, 5-3) = (1,1,0,0,1,1,2)\,. \qquad 3.1.10$$

To state the basic result of the Stanley Theory pf $\mathcal{P}$-partitions we need some notation. To begin, given a poset $\mathcal{P} = (\Omega, \preceq)$, the linear extensions of the partial order "$\preceq$" will be briefly referred to as the "*Standard Orders of* $\mathcal{P}$". If $\mathcal{P}$ has been given an injective labeling ω by the numbers $1, 2, \ldots, n$, then by reading its labels according to standard orders of $\mathcal{P}$ we obtain a collection of permutations $\sigma \in S_n$. Here and after we will call these permutations "ω-*Standard*" and we will denote their collection by "$ST_\omega(\mathcal{P})$". Finally, given a permutation $\sigma = \sigma_1\sigma_2\ldots\sigma_n$ a composition $p = (p_1, p_2, \ldots, p_n)$ will be called "σ-*compatible*" if its components satisfy the inequalities

$$p_n \geq 0 \quad \& \qquad \begin{cases} p_{r+1} \geq 1 & \text{if } \sigma_r > \sigma_{r+1} \\ \\ p_{r+1} \geq 0 & \text{otherwise} \end{cases} \qquad \text{for} \quad r = 1, 2, \ldots, n-1 \qquad 3.1.11$$

Since this condition essentially says that p majorizes the descent set of σ translated by 1 we will briefly express it by writing

$$\text{“} p >> 1 + D(\sigma) \text{”}$$

We now have the following fundamental fact.

Theorem 3.1.1

Let $\mathcal{P} = (\Omega, \preceq)$ be a poset with an injective labeling ω by the integers $1, 2, \ldots, n$. Then the map $f \to \big(\sigma(f), p(f)\big)$ defined by 3.1.6 and 3.1.7 is a bijection between the family $\mathcal{F}_{\mathcal{P},\omega}$ and the collection $\mathcal{C}(\mathcal{P}, \omega)$ of pairs (σ, p) where σ is ω-standard and p is σ-compatible. In symbols

$$\mathcal{C}(\mathcal{P}, \omega) \;=\; \Big\{ (\sigma, p) \,:\, \sigma \in ST_\omega(\mathcal{P}) \;\;\&\;\; p >> 1 + D(\sigma) \Big\} \qquad 3.1.12$$

Proof

For a given $f \in \mathcal{F}_\omega(\mathcal{P})$ let $\sigma(f) = \sigma_1\sigma_2\cdots\sigma_n$. Identifying the elements of Ω with their labels $1, 2, \ldots, n$, to show that $\sigma(f) \in ST_\omega(\mathcal{P})$ we need only verify that

$$\sigma_r \prec \sigma_s \quad \Rightarrow \quad r < s \qquad 3.1.13$$

Now recalling the construction that led to 3.1.6, we see that if $f(\sigma_r) = f(\sigma_s)$ then the definition in 3.1.2 yields that $\sigma_r < \sigma_s$. Moreover, σ_r and σ_s must lie in the same set A_i. But then σ_r must come before σ_s giving $r < s$ as desired. If $f(\sigma_r) \neq f(\sigma_s)$ then $\sigma_r \prec \sigma_s$ forces $f(\sigma_r) < f(\sigma_s)$ and this means that $\sigma_r \in A_i$ and $\sigma_s \in A_j$ with $i < j$, so we must again have $r < s$. This proves 3.1.13.

To show that $p(f)$ satisfies 3.1.11 note that, by its very construction, the descents of the permutation $\sigma(f)$ can only occur between two successive words "$\uparrow_\omega\ A_i$" and "$\uparrow_\omega\ A_{i+1}$". But if $\sigma_r \in A_i$ and $\sigma_{r+1} \in A_{i+1}$ then $f(\sigma_r) = v_i$ and $f(\sigma_{r+1}) = v_{i+1}$ give $p_{r+1}(f) = v_{i+1} - v_i \geq 1$ as desired.

Now the map $f \rightarrow \big(\sigma(f), p(f)\big)$ is clearly injective since we may simply recover f from the identity

$$f(\sigma_r) = p_1 + p_2 + \cdots + p_r \qquad 3.1.14$$

which reverses 3.1.7. To complete the proof we need only verify that this map is onto. Let then the pair $(\sigma, p) \in \mathcal{C}(\mathcal{P}, \omega)$ be given and let f be defined according to 3.1.14. We must show that $f \in \mathcal{F}_\omega(\mathcal{P})$ and that $(\sigma(f), p(f)) = (\sigma, p)$. To begin with, since σ is a linear extension of $\mathcal{P}$ we have that $\sigma_i \prec \sigma_j$ forces $i < j$ and thus the definition in 3.1.14 gives

$$f(\sigma_i) \leq f(\sigma_j)$$

as desired. Moreover, note that if $\sigma_i > \sigma_j$ then between i and j the permutation σ will necessarily have a descent and the σ-compatibility of p will force

$$f(\sigma_j) - f(\sigma_i) = p_j + p_{j-1} + \cdots + p_{i+1} > 0\ .$$

This shows that $f \in \mathcal{F}_\omega(\mathcal{P})$.

Finally, to construct the permutation $\sigma(f)$ according to the recipe in 3.1.6 we need to determine first the sets A_i. To this end let us decompose the permutation σ in the form

$$\sigma \;=\; B_1 B_2 \cdots B_h$$

where the B_j are the words obtained by cutting σ at its descents. Since these words are necessarily increasing, we may view their collection as a partition of the set $\{1, 2, \ldots, n\}$. Note then that, having constructed the sets A_i for the f defined by 3.1.14, we see that if $\sigma_r \in A_i$, then $p_{r+1} > 0$ will cause σ_{r+1} to be in A_{i+1} and this forces $A_1, A_2, \ldots, A_k$ to be a partition of $\{1, 2, \ldots, n\}$ which can be obtained by cutting the words B_j into successive segments. Putting it in another way, for some indices $1 \leq i_1 < i_2 < \cdots < i_{h-1} < k$ we will have

$$\begin{aligned}
&B_1 = \uparrow_\omega A_1 \uparrow_\omega A_2 \cdots \uparrow_\omega A_{i_1} \\
&B_2 = \uparrow_\omega A_{i_1+1} \uparrow_\omega A_{i_1+2} \cdots \uparrow_\omega A_{i_2} \\
&\cdots \\
&B_h = \uparrow_\omega A_{i_{h-1}+1} \uparrow_\omega A_{i_{h-1}+2} \cdots \uparrow_\omega A_k
\end{aligned}$$

But this gives

$$\sigma(f) \;=\; \uparrow_\omega A_1 \uparrow_\omega A_2 \cdots \uparrow_\omega A_k \;=\; B_1 B_2 \cdots B_h \;=\; \sigma$$

as desired. This given the identity

$$p(f) \;=\; p$$

immediately follows from the definition of f in 3.1.14. This completes our proof.

Theorem 3.1.1 yields a beautiful expansion for the polynomials $F_{\mathcal{P},\omega}$.

Theorem 3.1.2

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) = \sum_{\sigma \in ST_\omega(\mathcal{P})} \sum_{\substack{1 \le \beta_1 \le \beta_2 \le \cdots \le \beta_n \le N \\ \sigma_i > \sigma_{i+1} \Rightarrow \beta_i < \beta_{i+1}}} x_{\beta_1} x_{\beta_2} \cdots x_{\beta_n} \qquad 3.1.15$$

Proof

Using the map $f \to \big(\sigma(f), p(f)\big)$ from Theorem 3.1.1 and the definition in 3.1.3 we get that

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) = \sum_{(\sigma,p) \in \mathcal{C}(\mathcal{P},\omega)} x_{p_1} x_{p_1+p_2} \cdots x_{p_1+p_2+\cdots+p_n} \, .$$

However, 3.2.12 gives

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) = \sum_{\sigma \in ST_\omega(\mathcal{P})} \sum_{p >> 1 + D(\sigma)} x_{p_1} x_{p_1+p_2} \cdots x_{p_1+p_2+\cdots+p_n} \, . \qquad 3.1.16$$

Now $p >> 1 + D(\sigma)$ simply means that

$$\sigma_i > \sigma_{i+1} \implies p_1 + \cdots + p_i < p_1 + \cdots + p_{i+1}$$

and so we see that 3.1.15 is simply another way of writing 3.1.16.

Remark 3.1.1

We should mention that the inner sum in 3.1.15 is one of Gessel's "*Quasi-Symmetric*" functions []. To simplify some of our formulas, and to be consistent with the notation introduced in section 2.3, it will be good to represent these polynomials by a symbol indexed by a "strict" composition. (†) To this end if $p = (p_1, p_2, \ldots, p_k)$ and all $p_i \ge 1$ then we shall write

$$p \models n \;\; \longleftrightarrow p_1 + p_2 + \cdots + p_k = n$$

To such a composition p we shall associate the subset $S(p) \subseteq \{1, 2, \ldots, n\}$ defined by setting

$$S(p) \;=\; \{p_1, p_1 + p_2, \cdots, p_1 + p_2 + \cdots + p_{k-1}\} \qquad 3.1.17$$

This given, for $p \models n$ we shall here and after set

$$Q_p(x_1, x_2, \ldots, x_N) \;=\; \sum_{\substack{1 \le \beta_1 \le \beta_2 \le \cdots \le \beta_n \le N \\ i \in S(p) \Rightarrow \beta_i < \beta_{i+1}}} x_{\beta_1} x_{\beta_2} \cdots x_{\beta_n} \qquad 3.1.18$$

(†) That is an integral vector with all components ≥ 1

Now, recalling the definition in 2.3.17 of the composition $p(w)$ corresponding to the descent set of a word w, we see that the identity in 3.1.15 may be simply written as

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) = \sum_{\sigma \in ST_\omega(\mathcal{P})} Q_{p(\sigma)}(x_1, x_2, \ldots, x_N) \qquad 3.1.19$$

Remark 3.1.2

In view of the definition in 3.1.1, we see that for the example given in in 3.1.8 the labeling forces the elements of $\mathcal{F}_{\mathcal{P},\omega}(N)$ to be strictly increasing as we go NORTH-WEST and weakly increasing as we go NORTH-EAST. In particular, in this case, the family $\mathcal{F}_{\mathcal{P},\omega}$ can be identified with the collection of all column-strict tableaux of shape $(3,3,2)$. Recalling the definition of a Schur function S_λ as a sum of monomials corresponding to column-strict tableaux of shape λ we see that in this case we have

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) \;\; = \;\; S_{3,3,2}(x_1, x_2, \ldots, x_N)\,.$$

Clearly this is not an accident but a particular case of a general method for obtaining expansions of Schur functions in terms of quasi-symmetric functions. To state the result which follows, we need to make some notational conventions. Given a French skew diagram D with n cells and a standard filling τ of D we shall denote by $w(\tau)$ the permutation obtained by reading τ from from left to right, by rows starting from the top row. For instance if $D = 5442/311$ and

$$\tau = \begin{array}{ccccc} 2 & 10 & & & \\ & 3 & 7 & 9 & \\ & 1 & 4 & 8 & \\ & & & 5 & 6 \end{array} \qquad 3.1.20$$

then

$$w(\tau) = 2\,10\,3\,7\,9\,1\,4\,8\,5\,6 \qquad 3.1.21$$

Now we can construct from any skew diagram D a poset $\mathcal{P}_D$ by tilting the diagram 45^o counterclockwise and for two cells x, y set $x \prec y$ if and only if we can go from x to y by a sequence of NORTH-WEST and NORTH-EAST steps. In the display below we have illustrated the poset P_D corresponding to the shape $D = 5442/311$

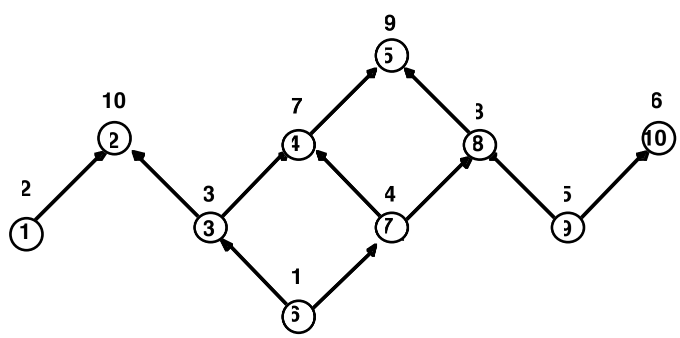


Fig. 3.1.22

In this display the numbers in circles are obtained by labeling the cells of D with $1, 2, \ldots, n = 10$ from left to right and from top to bottom. We shall here and after assume that the posets $\mathcal{P}_D$ are given an ω labeling constructed in this manner. This will be referred as the "*Natural Labeling of* $\mathcal{P}_D$". It should then be noted that every standard standard tableau τ of shape D will then give raise to a linear extension of $\mathcal{P}_D$. This should be quite clear from Fig. 3.1.22 where we have placed above the circles the corresponding entries of the tableau τ of 3.1.20. This given, to each standard tableau of shape D there will correspond an element $\sigma(\tau) \in ST_\omega(\mathcal{P}_D)$ obtained by reading the labels in the circles in the order given by the linear extension corresponding to τ. For instance in the case illustrated in Fig 3.1.22 we obtain the permutation

$$\sigma(\tau) \;=\; \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 1 & 3 & 7 & 9 & 10 & 4 & 8 & 5 & 2 \end{bmatrix} \qquad 3.1.23$$

We are now in a position to state and prove a basic expansion result for skew Schur functions.

Theorem 3.1.3

For any skew diagram D *we have*

$$S_D(x_1, x_2, \ldots, x_n) \;=\; \sum_{\tau \in ST(D)} Q_{p(\tau)}(x_1, x_2, \ldots, x_n) \qquad 3.1.24$$

Proof

Let $\mathcal{P}_D$ be the poset corresponding to D and let ω be the natural labeling of $\mathcal{P}_D$ obtained by the construction given above. We can easily see from the example displayed in 3.1.8 that the column strict tableaux of shape D may be identified with the ω-compatible $\mathcal{P}_D$-partitions. It thus follows from Theorem 3.1.2 that

$$S_D(x_1, x_2, \ldots, x_n) \;=\; \sum_{\sigma \in ST_\omega(\mathcal{P}_D)} Q_{p(\sigma)}(x_1, x_2, \ldots, x_n)\,.$$

Now from what we have observed it follows that this identity can be rewritten as

$$S_D(x_1, x_2, \ldots, x_n) \;=\; \sum_{\tau \in ST(D)} Q_{p(\sigma(\tau))}(x_1, x_2, \ldots, x_n)\,. \qquad 3.1.25$$

However, a glimpse at Fig. 3.1.22 and the permutation in 3.1.23 should reveal that the descents of $\sigma(\tau)$ occur precisely at the indices i of τ where $i+1$ is strongly NORTH and weakly WEST of i. But these are precisely the descents of τ itself. In other words by the notation we introduced in section 2.3 we have

$$p(\sigma(\tau)) \;=\; p(\tau)\,.$$

Substituting this in 3.1.25 gives 3.1.24 as desired, completing the proof.

Given two words $a = a_1 a_2 \cdots a_h$ and $b = b_1 b_2 \cdots b_k$, the collection of all words obtained by shuffling the letters of a and b (as if they were card decks) is called the "*shuffle of* a *and* b" and is denoted

$$\text{“} a \sqcup\!\sqcup b \text{”}.$$

For instance if $a = a_1 a_2$ and $b = b_1 b_2 b_3$ then

$$a \sqcup\!\sqcup b \;=\; \{a_1a_2b_1b_2b_3\,,\ a_1b_1a_2b_2b_3\,,\ a_1b_1b_2a_2b_3\,,\ a_1b_1b_2b_3a_2\,,\ b_1a_1a_2b_2b_3\,,$$
$$b_1a_1b_2a_2b_3\,,\ b_1a_1b_2b_3a_2\,,\ b_1b_2a_1a_2b_3\,,\ b_1b_2a_1b_3a_2\,,\ b_1b_2b_3a_1a_2\,\}$$

The following result shows the peculiar way by which quasisymmetric functions multiply.

Theorem 3.1.4

For $\alpha = \alpha_1\alpha_2\cdots\alpha_h \in S_h$ and $\beta = \beta_1\beta_2\cdots\beta_k \in S_k$ we have (for sufficiently large N)

$$Q_{p(\alpha)}(x_1, x_2, \ldots, x_N)Q_{p(\beta)}(x_1, x_2, \ldots, x_N) = \sum_{\sigma\in\alpha \sqcup\!\sqcup 1_h\otimes\beta} Q_{p(\sigma)}(x_1, x_2, \ldots, x_N) \qquad 3.1.26$$

Proof

Let $\mathcal{P}_h$ denote the ordinary chain

$$\mathcal{P}_h = (\{1, 2, \ldots, h\}, \leq)\,.$$

and let us label the elements $1, 2, \ldots, h$ by $\alpha_1, \alpha_2, \ldots, \alpha_h$ respectively. Since $\mathcal{P}_h$ is linearly ordered the collection $ST_\alpha(\mathcal{P}_h)$ reduces to the single permutation α. Thus from 3.1.19 we derive that

$$F_{\mathcal{P}_h,\alpha}(x_1, x_2, \ldots, x_N) \;=\; Q_{p(\alpha)}(x_1, x_2, \ldots, x_N)\,. \qquad 3.1.27$$

Similarly, if $\mathcal{P}_k = (\{1, 2, \ldots, k\}, \leq)$ and we label its elements $\beta_1, \beta_2, \ldots, \beta_k$ we get

$$F_{\mathcal{P}_k,\beta}(x_1, x_2, \ldots, x_N) \;=\; Q_{p(\beta)}(x_1, x_2, \ldots, x_N)\,. \qquad 3.1.28$$

Now let $\mathcal{P} = \mathcal{P}_h \cup \mathcal{P}_k$ be the poset consisting of the simple disjoint union of these two chains and let ω be the labeling of $\mathcal{P}$ obtained by giving the elements of $\mathcal{P}_h$ the labels $\alpha_1, \alpha_2, \ldots, \alpha_h$ and the elements of $\mathcal{P}_k$ the labels

$$h + \beta_1\;,\;\, h + \beta_2\;,\;\, \ldots\;,\;\, h + \beta_k\,. \qquad 3.1.29$$

This given it is easy to see that every ω-compatible $\mathcal{P}$-partition $f \in \mathcal{F}_{\mathcal{P},\omega}$ is simply obtained by choosing a pair $f_1 \in \mathcal{F}_{\mathcal{P}_h,\alpha}(N)$ and $f_2 \in \mathcal{F}_{\mathcal{P}_k,\beta}(N)$ and transplanting them onto the $\mathcal{P}_h$ and $\mathcal{P}_k$ portions of $\mathcal{P}$. In fact, the ω-compatibility of f_1 is trivial and that of f_2 follows from the fact that the labeling in 3.1.29 has the same descent set as the labeling $\beta_1, \beta_2, \ldots, \beta_k$. Thus it follows that in this case

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) \;=\; Q_{p(\alpha)}(x_1, x_2, \ldots, x_N)\, Q_{p(\beta)}(x_1, x_2, \ldots, x_N)\,. \qquad 3.1.30$$

Now it turns out that the desired identity in 3.1.26 is obtained by computing the same polynomial by means of formula 3.1.19. In fact, it is easy to see that here the elements of $ST_\omega(\mathcal{P})$ are none other than the shuffles of $\alpha_1, \alpha_2, \ldots, \alpha_h$ with the labels in 3.1.29. In our notation these are simply the permutations in

$$\alpha \;\sqcup\!\sqcup\; 1_h \otimes \beta$$

Thus in this case 3.1.19 may be rewritten as

$$F_{\mathcal{P},\omega}(x_1, x_2, \ldots, x_N) \;=\; \sum_{\sigma\in\alpha \sqcup\!\sqcup 1_h\otimes\beta} Q_{p(\sigma)}(x_1, x_2, \ldots, x_N)\,.$$

This completes our argument.

3.2. The Stanley Symmetric Function of a Permutation.

Early in the summer of 1982 Richard Stanley started an investigation aimed at the enumeration of reduced decompositions. This was prompted by his discovery that data gathered in previous work [] showed that the number of reduced decomposition of the the top element of S_n for $n = 2, 3, 4, 5, 6$ is equal to the number of standard tableaux of corresponding staircase shape. Given his previous work, in particular formula 3.1.19, he was led to the bold step of setting for any given $\sigma \in S_n$ and $N > l(\sigma)$

$$F_\sigma(x_1, x_2, \ldots, x_N) \;=\; \sum_{w \in RED(\sigma)} Q_{p(w)}(x_1, x_2, \ldots, x_N)\,. \qquad 3.2.1$$

Unbeknown to him at the time, he was essentially discovering a natural generalization of "*Skew Schur Functions*". Experimentations with examples that can be obtained by hand computations led him to conjecture that F_σ is a Symmetric Function with a Schur Function expansion of the form

$$F_\sigma(x_1, x_2, \ldots, x_N) \;=\; \sum_{\lambda \in \mathcal{C}(\sigma)} a_\lambda(\sigma)\, S_\lambda(x_1, x_2, \ldots, x_N)\;. \qquad 3.2.2$$

with $\mathcal{C}(\sigma)$ a suitable collection of shapes and the $a_\lambda(\sigma)$ certain positive integers. His investigations led him to a seminal publication [] where he presented a number of results supporting his conjectures. In particular he proved the symmetry and showed the containement

$$\mathcal{C}(\sigma) \subseteq \big\{ \lambda \,:\, \lambda(\sigma^{-1}) \le \lambda \le \lambda'(\sigma) \big\}\,. \qquad 3.2.3$$

In particular he derived (see 2.3.12 and Theorem 3.2.3 below) that for σ vexillary

$$F_\sigma(x_1, x_2, \ldots, x_N) \;=\; S_{\lambda'(\sigma)}(x_1, x_2, \ldots, x_N)\,. \qquad 3.2.4$$

This allowed him to completely settle the case of the top element of S_n. However he was not able to prove Schur positivity (i.e. $a_\lambda(\sigma) > 0$ in 3.2.2) nor identify the collection $\mathcal{C}(\sigma)$. In subsequent years all of his conjectures were proved and even some analogous results were established for other Coxeter groups, in a variety of papers [],[],[]. The methods used ranged from purely combinatorial, to representation theoretical and algebraic geometrical. In reviewing this literature we discovered that a relatively simple and very accessible proof of the Schur positivity of F_σ can be obtained by suitably combining a number of results from a variety of sources. To be precise, note that as a corollary of Theorem 2.4.4 we obtain the following remarkably beautiful solution of the Schur positivity problem for F_σ.

Theorem 3.2.1

On the validity of Theorem 2.4.1, for any permutation σ we have

$$F_\sigma(x_1, x_2, \ldots, x_N) \;=\; \sum_{\sigma' \in LeavesLS(\sigma)} S_{\lambda'(\sigma')}(x_1, x_2, \ldots, x_N) \qquad 3.2.5$$

In particular, for the collection of shapes occurring in 3.2.2 we obtain that

$$\mathcal{C}(\sigma) \;=\; \big\{ \lambda \,:\, \lambda = \lambda'(\sigma') \;\; \text{for some} \;\; \sigma' \in LeavesLS(\sigma) \big\} \qquad 3.2.6$$

Moreover, form 3.2.5 we derive that the multiplicities $a_\lambda(\sigma)$ have a very simple combinatorial interpretation, namely

$$a_\lambda(\sigma) \;=\; \#\big\{ \sigma' \in LeavesLS(\sigma) \,:\, \lambda'(\sigma') = \lambda \,\big\} \qquad 3.2.7$$

Proof

From the definitions of $\Xi(\sigma)$ and $\Sigma(\lambda)$ given in 2.3.19 we see that

$$F_\sigma(x_1,x_2,\ldots,x_n) \;=\; \Xi(\sigma)\Big|_{x_p\to Q_p(x_1,x_2,\ldots,x_n)}\,. \qquad 3.2.8$$

Now using 2.4.36 we derive that

$$F_\sigma(x_1,x_2,\ldots,x_n) \;=\; \sum_{\sigma'\in Leaves LS(\sigma)} \Sigma\big(\lambda'(\sigma')\big)\Big|_{x_p\to Q_p(x_1,x_2,\ldots,x_n)}\,. \qquad 3.2.9$$

Now note that from the definition in 2.3.21 and Theorem 3.1.3 we get that for any partition λ we have

$$S_\lambda(x_1,x_2,\ldots,x_n) \;=\; \Sigma(\lambda)\Big|_{x_p\to Q_p(x_1,x_2,\ldots,x_n)}\,. \qquad 3.2.10$$

This given we see that 3.2.5 follows from 3.2.9. This completes the proof since 3.2.6 1nd 3.2.7 are immediate consequences of 3.2.5.

Remark 3.2.1

It should be noted that also the containement in 3.2.3 follows from Theorem 3.2.1. Indeed we see from Lemma 2.4.1 that for every regular child σ' of a permutation σ we have

$$\lambda(\sigma)\le\lambda(\sigma')\le\lambda'({\sigma'}^{-1})\le\lambda'(\sigma^{-1})$$

and since $\lambda(\sigma) = \lambda(1\times\sigma)$ we see that these inequalities must hold also for a lateral child. Applying them recursively yields that they will have to hold as well for any leaf σ' of the LS tree of σ. Thus 3.2.3 follows from 3.2.6.

Our proof of Theorem 2.4.1, on which the validity of Theorem 3.2.1 depends, will be given in the next section. It will be based on the Theory of Schubert polynomials together with some of the identities proved in [], [] and []. In the remainder of this section we shall present some results and proofs given in [] and []. In particular we shall include here the very beautiful argument given by Fomin and Stanley in [] proving the symmetry of F_σ. Of course, also this symmetry is a consequence of 3.2.5. However, even though most of what we ever wanted to show follows from Theorem 3.2.1, there are a number of beautiful arguments and results in this theory that are worth relating. So it will be worthwhile to include some of them here, even at the expense of ending up with more than one proof of the same result.

In [] Fomin and Stanley base their arguments on the so called "*Nil-Coxeter*" algebra $\mathcal{NC}_n$. Using this device they were not only able to prove the symmetry of F_σ but also could derive in a very efficient way some of the basic properties of Schubert polynomials. This given it will be most appropriate to introduce it in this section. The definition of $\mathcal{NC}_n$ is quite immediate. It is simply a K-algebra with generators

$$u_1,u_2,\ldots,u_{n-1}\,, \qquad 3.2.11$$

together with an identity "1", and relations

$$\begin{array}{ll} a) & u_i^2=0\,,\\ b) & u_i\,u_j=u_j\,u_i \quad \text{when } |i-j|>1\,,\\ c) & u_i\,u_{i+1}\,u_i=u_{i+1}\,u_i\,u_{i+1} \quad \text{for } 1\le i\le n-2\,. \end{array} \qquad 3.2.12$$

Here K needs only be the ring of polynomials with integer coefficients in the variables $x_1, x_2, \dots x_N, x, y$. We shall see that such an algebra has a natural faithful representation in terms of the Lascoux-Schützenberger divided difference operators δ_i introduced in the next section.

The relations in 3.2.12 assure that for any word $w = a_1 a_2 \cdots a_l$ we shall have

$$u_{a_1} u_{a_2} \cdots u_{a_l} \neq 0$$

if and only if w is a reduced word of some permutation σ. Moreover, using $b)$ and $c)$ we can show that if $w = a_1 a_2 \cdots a_l$ and $w' = a'_1 a'_2 \cdots a'_l$ are both reduced words for the same permutation σ then we necessarily have

$$u_{a_1} u_{a_2} \cdots u_{a_l} \;\; = \;\; u_{a'_1} u_{a'_2} \cdots u_{a'_l} \; .$$

This means that to any $\sigma \in S_n$ we can associate a well defined element $u_\sigma \in \mathcal{NC}_n$ simply by setting for any reduced word $w = a_1 a_2 \cdots a_l \in RED(\sigma)$

$$u_\sigma \;\; = \;\; u_{a_1} u_{a_2} \cdots u_{a_l} \; . \qquad 3.2.13$$

This given, Fomin and Stanley set

$$\mathcal{A}_i(x) = (1 + x\, u_{n-1})(1 + x\, u_{n-2}) \cdots (1 + x\, u_i) \; . \qquad 3.2.14$$

and obtain the following basic commutativity relations.

Proposition 3.2.1

$$\mathcal{A}_i(x)\, \mathcal{A}_i(y) \;\; = \;\; \mathcal{A}_i(y)\, \mathcal{A}_i(x) \qquad (\text{ for } i = 1, 2, \dots, n-1 \text{ }) \qquad 3.2.14$$

Proof

Note that for $i = n-1$ the identity in 3.2.14 reduces to

$$(1 + x\, u_{n-1})(1 + y\, u_{n-1}) \;\; = \;\; (1 + y\, u_{n-1})(1 + x\, u_{n-1}) \, . \qquad 3.2.15$$

This is trivially true since setting

$$h_i(x) = (1 + x\, u_i) \qquad 3.2.16$$

from $a)$ of 3.2.12 we derive that

$$h_i(x) h_i(y) \;\; = \;\; h_i(x+y) \;\; = \;\; h_i(y) h_i(x) \, . \qquad 3.2.17$$

So the idea is to prove 3.2.14 by descent induction on i. Now the crucial identity here is a beautiful extension of 3.2.12 $c)$, namely

$$h_i(x) h_{i+1}(x+y) h_i(y) \;\; = \;\; h_{i+1}(y) h_i(x+y) h_{i+1}(x) \, . \qquad 3.2.18$$

This can be easily verified by means of $a)$ and $c)$ of 3.2.12. Now, assume that we have shown

$$\mathcal{A}_{i+1}(x)\, \mathcal{A}_{i+1}(y) \;\; = \;\; \mathcal{A}_{i+1}(y)\, \mathcal{A}_{i+1}(x) \qquad 3.2.19$$

This given, we have

$$\begin{array}{rcl}
\mathcal{A}_i(x)\mathcal{A}_i(y) & = & \mathcal{A}_{i+1}(x)h_i(x)\mathcal{A}_{i+2}(y)h_{i+1}(y)h_i(y) \\
(\text{ using 3.2.12 } b)\ ) & = & \mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)h_i(x)h_{i+1}(y)h_i(y) \\
(\text{ using 3.2.17 }) & = & \mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)h_i(x)h_{i+1}(y)h_i(y-x)h_i(x) \\
(\text{ using 3.2.18 }) & = & \mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)h_{i+1}(y-x)h_i(y)h_{i+1}(x)h_i(x) \\
(\text{ using 3.2.17 }) & = & \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y)h_{i+1}(-x)h_i(y)h_{i+1}(x)h_i(x) \\
(\text{ using 3.2.19 }) & = & \mathcal{A}_{i+1}(y)\mathcal{A}_{i+1}(x)h_{i+1}(-x)h_i(y)h_{i+1}(x)h_i(x) \\
(\text{ using 3.2.17 }) & = & \mathcal{A}_{i+1}(y)\mathcal{A}_{i+2}(x)h_i(y)h_{i+1}(x)h_i(x) \\
(\text{ using 3.2.12 } b)\ ) & = & \mathcal{A}_{i+1}(y)h_i(y)\mathcal{A}_{i+2}(x)h_{i+1}(x)h_i(x) \ = \ \mathcal{A}_i(y)\mathcal{A}_i(x)
\end{array}$$

completing the induction and the proof of the proposition.

The relevance of these computations in our context stems from the following remarkable identities

Proposition 3.2.2

Given a permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ set

$$\sigma^* = \sigma_1^*\sigma_2^*\cdots\sigma_n^* \tag{3.2.20}$$

with

$$\sigma_i^* \ = \ n+1-\sigma_{n+1-i} \qquad (\text{ for } i=1,2,\ldots,n\ ) \tag{3.2.21}$$

Then for $N > l = l(\sigma)$ we have

$$F_{\sigma^*}(x_1,x_2,\ldots,x_n) = \sum_{a_1a_2\cdots a_l \in RED(\sigma)} \ \sum_{\substack{1\leq\beta_1\leq\beta_2\leq\cdots\leq\beta_l\leq N \\ a_i<a_{i+1} \Rightarrow \beta_i<\beta_{i+1}}} x_{\beta_1}x_{\beta_2}\cdots x_{\beta_l} \tag{3.2.22}$$

in particular

$$F_{\sigma^*}(x_1,x_2,\ldots,x_n) \ = \ \mathcal{A}_1(x_1)\mathcal{A}_1(x_2)\cdots\mathcal{A}_1(x_N)\Big|_{u_\sigma} \tag{3.2.23}$$

Proof

Note first that if we turn upside down the line diagram of a reduced decomposition $a_1a_2\cdots a_l$ and replace each label "i" by the label "$n+1-i$" the result will simply be the line diagram of $n-a_1n-a_2\cdots n-a_l$. Since this replacement changes the target permutation σ into σ^* we deduce that we have

$$a_1a_2\cdots a_l \in RED(\sigma) \quad \Longleftrightarrow \quad n-a_1n-a_2\cdots n-a_l \in RED(\sigma^*)\ . \tag{3.2.24}$$

This means that if

$$w = a_1a_2\cdots a_l \qquad \text{and} \qquad w^* = n-a_1n-a_2\cdots n-a_l$$

Then the descent sets of w and w^* are complements of one another. In symbols

$$D(w^*) \ = \ {}^cD(w) = \{1,2,\ldots,l-1\} - D(w) \tag{3.2.25}$$

Now the definition in 3.2.1 may also be written as

$$F_\sigma(x_1,x_2,\ldots,x_n) = \sum_{a_1a_2\cdots a_l\in RED(\sigma)} \sum_{\substack{1\le\beta_1\le\beta_2\le\cdots\le\beta_l\le N \\ a_i>a_{i+1}\ \Rightarrow\ \beta_i<\beta_{i+1}}} x_{\beta_1}x_{\beta_2}\cdots x_{\beta_l} \qquad 3.2.26$$

In particular, using 3.2.24 we derive that for F_{σ^*} we have the expansion

$$F_{\sigma^*}(x_1,x_2,\ldots,x_n) = \sum_{a_1a_2\cdots a_l\in RED(\sigma)} \sum_{\substack{1\le\beta_1\le\beta_2\le\cdots\le\beta_l\le N \\ n-a_i>n-a_{i+1}\ \Rightarrow\ \beta_i<\beta_{i+1}}} x_{\beta_1}x_{\beta_2}\cdots x_{\beta_l}$$

and this is simply another way of writing 3.2.22.

Finally, note that when we expand the product in the right hand side of 3.2.23, we obtain terms of the form

$$x_{\beta_1}x_{\beta_2}\cdots x_{\beta_m}\ u_{a_1}u_{a_2}\cdots u_{a_m}\Big|_{u_\sigma}$$

with

$$\beta_1\le\beta_2\le\cdots\le\beta_m$$

satisfying

$$a_i<a_{i+1} \quad\Longrightarrow\quad \beta_i<\beta_{i+1}\ .$$

This is because from the definition in 3.2.14 we get that two successive factors $x_{\beta_i}u_{a_i}$ and $x_{\beta_{i+1}}u_{a_{i+1}}$ with $\beta_i = \beta_{i+1}=r$ coming form the same $\mathcal{A}_r(x_r)$ in 3.2.23 will necessarily also have $a_i>a_{i+1}$.

Now because of 3.2.12 a) the only terms that survive are those for which $m=l$,

$$u_{a_1}u_{a_2}\cdots u_{a_l} \ = \ u_\sigma\ .$$

and

$$a_1a_2\cdots a_l\in RED(\sigma)\ .$$

Thus 3.2.23 follows from 3.2.22. This completes our argument.

As a corollary of Proposition 3.2.2 we obtain

Theorem 3.2.2

For any permutation σ the Stanley polynomial $F_\sigma(x_1,x_2,\ldots,x_n)$ is a symmetric fumction of $x_1,x_2,\ldots,x_n$

Proof

From 3.2.23 we derive that

$$F_\sigma(x_1,x_2,\ldots,x_n) \ = \ \mathcal{A}_1(x_1)\mathcal{A}_1(x_2)\cdots\mathcal{A}_1(x_N)\Big|_{u_{\sigma^*}}\ . \qquad 3.2.27$$

Thus the assertion is a simple consequence of Proposition 3.2.1.

Stanley's proof of the inclusion in 3.2.3 is based on the following two auxiliary results.

Proposition 3.2.2

$$\mathcal{C}(\sigma) \subseteq \Big\{ \lambda \,:\, \lambda \subseteq \lambda'(\sigma) \Big\} \qquad 3.2.28$$

Proof

Since we have proved that F_σ is symmetric we shall have an expansion of the form

$$F_\sigma(x_1, x_2, \ldots, x_n) \;=\; \sum_\mu b_\mu(\sigma)\, m_\mu(x_1, x_2, \ldots, x_n) \qquad 3.2.29$$

where "m_μ" denotes the monomial symmetric funtion corresponding to μ and the $b_\mu(\sigma)$ are suitable non-negative integer coefficients. In view of the expansion in 3.2.26 we see that $b_\mu(\sigma) > 0$ if and only if at least one of the summands in 3.2.26 yields the leading monomial of m_μ. In other words, if $b_\mu(\sigma) > 0$ for

$$\mu = (\, \mu_1 \geq \mu_2 \geq \cdots \geq \mu_k > 0\,) \vdash l$$

then from some word $w = a_1 a_2 \cdots a_l \in RED(\sigma)$ we have

$$x_{\beta_1} x_{\beta_2} \cdots x_{\beta_l} \;=\; x_1^{\mu_1} x_2^{\mu_2} \cdots x_k^{\mu_k}$$

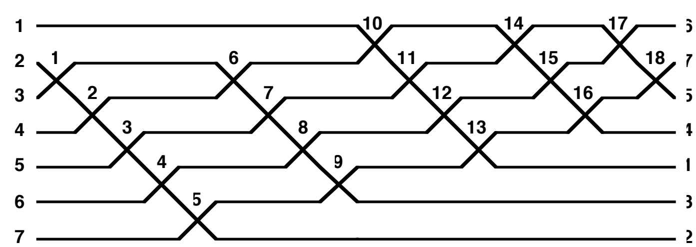
with $\beta_1 \leq \beta_2 \cdots \leq \beta_l$ and $\beta_i < \beta_{i+1}$ when $a_i > a_{i+1}$. Now this implies that the descents of w must be all contained in the set

$$\big\{\mu_1 \,,\, \mu_1 + \mu_2 \,,\, \mu_1 + \mu_2 + \mu_3 \,,\, \ldots \,,\, \mu_1 + \mu_2 + \cdots + \mu_{k-1} \big\}\,.$$

Equivalently, we must have the inequalities

$$a_1 < a_2 < \cdots < a_{\mu_1} \;,\; a_{\mu_1+1} < a_{\mu_1+2} < \cdots < a_{\mu_1+\mu_2} \;,\; \cdots \;,\; a_{\mu_1+\cdots+\mu_{k-1}+1} < a_{\mu_1+\cdots+\mu_{k-1}+2} < \cdots < a_l$$

To see what this tells us about the circle diagram of σ we only need to have a look at the corresponding line diagram $\mathcal{M}(a_1 a_2 \cdots a_l)$. To this end we have depicted below the case $w = 23456 \cdot 2345 \cdot 1234 \cdot 123 \cdot 12$ and $\mu = (5, 4, 4, 3, 2)$,



Let us imagine that we break up the construction of our diagram into k stages containing $\mu_1, \mu_2, \ldots, \mu_k$ steps respectively. In this case we obtain the successsion of diagrams

$$\mathcal{M}(a_1, a_2, \ldots, a_5) \to \mathcal{M}(a_1, a_2, \ldots, a_9) \to \mathcal{M}(a_1, a_2, \ldots, a_{13}) \to \mathcal{M}(a_1, a_2, \ldots, a_{16}) \to \mathcal{M}(a_1, a_2, \ldots, a_{18})$$

Now recall that, according to definition, 2.1 an "$\times$" at the k^{th} step contributes a circle labelled "k" in position (i, j) of $CD(\sigma)$ if that "$\times$" interchanges the i-line with the σ_j-line. In this particular example, the first stage creates

5 labelled circles. Due to the fact that $a_1 < a_2 < \cdots < a_5$ our definition implies that these circles will fall in 5 different columns. Proceeding with our construction, in the second stage we add 4 more circles some of which could land in the same column as the ones created in the first stage, but due to the fact that $a_6 < a_7 < \cdots < a_9$ they themselves will fall in 4 different columns. Similarly in the third stage we add 4 more circles in 4 different columns. Here some of these circles could land in the same column as one or two circles created in the two previous stages.

In the general case after $r \leq k$ stages we will have created

$$\mu_1 + \mu_2 + \cdots + \mu_r \tag{3.2.30}$$

labelled circles and, due to the fact that during each stage the a_i increase the circles created within a stage will land in separate columns. This causes circles appearing in the same column to come from different stages. Consequently, after r stages there will be at most r circles in any given column. This means that if we push these circles up along their column until they are tightly packed, they will necessarily fall in the first r lines of the circle diagram. On the other hand, if, after we finish the construction, we tightly pack all the circles of $CD(\sigma)$ in the same manner, we see from the defintion 2.3.3 of the code of σ, that the number of circles that will be packed in the first r rows is given by the expression

$$\sum_{i=1}^{n} c_i(\sigma) \wedge r\,,$$

where $a \wedge b = min(a, b)$. But since the shape (see definition 2.3.1) is only a rearrangement of the code we necessarily have the equalities

$$\sum_{i=1}^{n} c_i(\sigma) \wedge r\,, \;=\; \sum_{i} \lambda_i(\sigma) \wedge r \;=\; \lambda'_1(\sigma) + \lambda'_2(\sigma) + \cdots + \lambda'_r(\sigma)\;. \tag{3.2.31}$$

Since in the process of constructing the corresponding sequence of balanced tableaux

$$T(a_1 a_2 \cdots a_{\mu_1}) \longrightarrow T(a_1 a_2 \cdots a_{\mu_2}) \longrightarrow \cdots \longrightarrow T(a_1 a_2 \cdots a_l)$$

pairs of circles in different columns remain in different columns and pairs of cicles in the same column remain in the same column, it follows that the circles counted by 3.2.30 will be a subset of those counted by 3.2.31, and thus we must necessarily have

$$\mu_1 + \mu_2 + \cdots + \mu_r \leq \lambda'_1(\sigma) + \lambda'_2(\sigma) + \cdots + \lambda'_r(\sigma)\,.$$

In summary we have shown that

$$b_\mu(\sigma) > 0 \qquad \Longrightarrow \qquad \mu \leq \lambda'(\sigma)\,.$$

Thus the expansion in 3.2.29 may be rewritten as

$$F_\sigma(x_1, x_2, \ldots, x_n) \;=\; \sum_{\mu \leq \lambda'(\sigma)} b_\mu(\sigma)\; m_\mu(x_1, x_2, \ldots, x_n)\,. \tag{3.2.32}$$

Now recall from Symmetric Function Theory that the "monomial" and "Schur" bases are related by upper unitriangular matrices. Thus we may write

$$m_\mu(x_1, x_2, \ldots, x_n) \;=\; \sum_{\lambda \le \mu} S_\lambda(x_1, x_2, \ldots, x_n)\, H_{\lambda\mu}\,. \tag{3.2.33}$$

Substituting this in 3.2.32 gives

$$\begin{aligned} F_\sigma(x_1, x_2, \ldots, x_n) &= \sum_{\mu \le \lambda'(\sigma)} b_\mu(\sigma) \sum_{\lambda \le \mu} S_\lambda(x_1, x_2, \ldots, x_n)\, H_{\lambda\mu}\,. \\ &= \sum_{\lambda} S_\lambda(x_1, x_2, \ldots, x_n) \sum_{\lambda \le \mu \le \lambda'(\sigma)} b_\mu(\sigma)\;\; H_{\lambda\mu}\,. \end{aligned}$$

This shows that the coefficients $a_\lambda(\sigma)$ in the expansion 3.2.2 satisfy

$$a_\lambda(\sigma) \;=\; \begin{cases} \sum_{\lambda \le \mu \le \lambda'(\sigma)} b_\mu(\sigma)\;\; H_{\lambda\mu} & \text{if } \lambda \le \lambda'(\sigma)\,, \\ 0 & \text{otherwise.} \end{cases}$$

This proves 3.2.28 and completes our proof.

We now need two further properties of the permutation σ^*.

Proposition 3.2.3

For any permutation σ we have for $N \ge l(\sigma)$

$$\lambda(\sigma^*) \;=\; \lambda(\sigma^{-1}) \tag{3.2.34}$$

$$F_\sigma(x_1, x_2, \ldots, x_n) \;=\; \omega F_{\sigma^*}(x_1, x_2, \ldots, x_n) \tag{3.2.35}$$

where ω denotes the fundamental symmetric function involution.

Proof

Note that from 3.2.21 we get that the code of σ^* is given by the equalities

$$\begin{aligned} c_i(\sigma^*) &= \#\Big\{\, j > i \,:\, n+1-\sigma_{n+1-i} \;>\; n+1-\sigma_{n+1-j} \,\Big\} \\ &= \#\Big\{\, j > i \,:\, \sigma_{n+1-j} > \sigma_{n+1-i} \,\Big\}\,. \end{aligned}$$

Now this may be rewritten as

$$\begin{aligned} c_{n+1-i}(\sigma^*) &= \#\Big\{\, n+1-j > n+1-i \,:\;\; \sigma_j \;>\; \sigma_i \,\Big\} \\ &= \#\Big\{\, j < i \,:\, \sigma_j \;> \sigma_i \,\Big\}\,. \end{aligned}$$

This proves 3.2.34 since

$$\#\Big\{\, j < i \,:\, \sigma_j \;> \sigma_i \,\Big\} \;=\; c_i(\sigma^{-1})\,.$$

To prove 3.2.35 note that from 3.2.25 and 3.2.1 it follows that we may write for $l(\sigma) = l$

$$F_{\sigma^*}(x_1, x_2, \ldots, x_n) = \sum_{w \in RED(\sigma)} Q_{p(\ ^cD(w),l)}(x_1, x_2, \ldots, x_n) \, .$$

Now using 3.2.2 and 3.1.24 this may also be rewritten as

$$F_{\sigma^*}(x_1, x_2, \ldots, x_n) = \sum_{\lambda \in \mathcal{C}(\sigma)} a_\lambda(\sigma) \sum_{\tau \in ST(\lambda)} Q_{p(\ ^cD(\tau),l)}(x_1, x_2, \ldots, x_n) \qquad 3.2.36$$

On the other hand, since transposing a standard tableau complements its descent set, again from 3.1.24 we obtain that for any partition $\lambda \vdash l$ we have (for $N \geq l$)

$$\sum_{\tau \in ST(\lambda)} Q_{p(\ ^cD(\tau),l)}(x_1, x_2, \ldots, x_n) = S_{\lambda'}(x_1, x_2, \ldots, x_n) = \omega S_\lambda(x_1, x_2, \ldots, x_n) \, .$$

Substituting this in 3.2.36 gives 3.2.35 precisely as asserted.

We now have all we need to give Stanley's proof of 3.2.3. More precisely he obtains.

Theorem 3.2.3

For any permutation σ we have

$$F_\sigma(x_1, x_2, \ldots, x_n) = \sum_{\lambda(\sigma^{-1}) \leq \lambda \leq \lambda'(\sigma)} a_\lambda(\sigma) S_\lambda(x_1, x_2, \ldots, x_n) \qquad 3.2.37$$

with

$$a) \quad a_{\lambda(\sigma)}(\sigma^{-1}) = 1 \quad \text{and} \quad b) \quad a_{\lambda'(\sigma)}(\sigma) = 1 \qquad 3.2.38$$

Proof

Applying 3.2.28 to σ^* we can write

$$F_{\sigma^*}(x_1, x_2, \ldots, x_n) = \sum_{\lambda \leq \lambda'(\sigma^*)} a_{\lambda(\sigma^*)} S_\lambda(x_1, x_2, \ldots, x_n) \, ,$$

and from 3.2.35 we get that

$$F_\sigma(x_1, x_2, \ldots, x_n) = \sum_{\lambda \leq \lambda'(\sigma^*)} a_{\lambda(\sigma^*)} S_{\lambda'}(x_1, x_2, \ldots, x_n) \, .$$

Changing variable of summation from λ to λ' yields that this may also be rewritten as

$$F_\sigma(x_1, x_2, \ldots, x_n) = \sum_{\lambda' \leq \lambda'(\sigma^*)} a_{\lambda'(\sigma^*)} S_\lambda(x_1, x_2, \ldots, x_n) \, ,$$

and using 3.2.34 together with the fact that conjugating reverses dominance we finally get that

$$F_\sigma(x_1, x_2, \ldots, x_n) = \sum_{\lambda \geq \lambda(\sigma^{-1})} a_{\lambda'(\sigma^*)} S_\lambda(x_1, x_2, \ldots, x_n) \, .$$

Since Proposition 3.2.2 gives also

$$F_\sigma(x_1, x_2, \ldots, x_n) \;=\; \sum_{\lambda \leq \lambda'(\sigma)} a_{\lambda(\sigma)} S_\lambda(x_1, x_2, \ldots, x_n)\,,$$

we see that 3.2.37 must necessarily hold true as well.

It is easily seen from the above argument that 3.2.38 a) for σ implies 3.2.38 b) for σ^*. Thus we need to establish only one of these equalities. We shall prove 3.2.38 b). To this end we must show that in the expansion 3.2.26 there is one and only way to obtain the equality

$$x_{\beta_1} x_{\beta_2} \cdots x_{\beta_l} \;=\; x_1^{\mu_1} x_2^{\mu_2} \cdots x_k^{\mu_k}\,, \qquad 3.2.39$$

when

$$\mu_i = \lambda_i'(\sigma) \qquad (\text{ for } i = 1, 2, \ldots, k\ ) \qquad 3.2.40$$

and k is the number of parts of $\lambda'(\sigma)$. This implies that in the expansion 3.2.32 we must have $b_{\lambda_1'(\sigma)} = 1$ and then 3.2.38 follows since $H_{\mu\mu} = 1$ in 3.2.33.

As we noted in the proof of Proposition 3.2.2, we may have the equality in 3.2.39 only if the associated reduced word $w = a_1 a_2 \cdots a_l$ satisfies the inequalities

$$a_1 < a_2 < \cdots < a_{\mu_1}\ ,\ a_{\mu_1+1} < a_{\mu_1+2} < \cdots < a_{\mu_1+\mu_2}\ ,\ \cdots\ ,\ a_{\mu_1+\cdots+\mu_{k-1}+1} < a_{\mu_1+\cdots+\mu_{k-1}+2} < \cdots < a_l \qquad 3.2.41$$

To see that 3.2.40 and 3.2.41 determine the a_i uniquely we need only make one fundamental observation. Namely that in any column of a line diagram one "*high*" label gets interchanged with a "*low*" label.

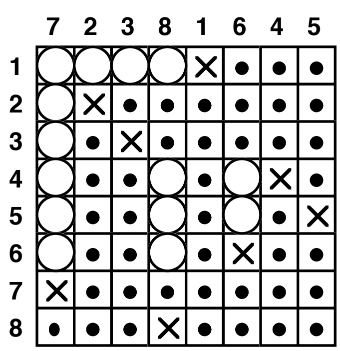
Now if we construct the line diagram $\mathcal{M}(a_1, a_2, \cdots, a_l)$ in stages

$$\cdots \longrightarrow \mathcal{M}(a_1, a_2, \cdots, a_{\mu_1+\cdots+\mu_{r-1}}) \longrightarrow \mathcal{M}(a_1, a_2, \cdots, a_{\mu_1+\cdots+\mu_r}) \longrightarrow \cdots$$

for $r = 2, 3, \ldots, k$, it follows that at the r^{th} stage exactly μ_r distinct high labels are interchanged with μ_r low labels (not necessarily distinct). We claim that the requirements in 3.2.40 and 3.2.41 force the high labels involved at the r^{th} stage to be the collection

$$M_{\geq r}(\sigma) = \big\{\sigma_i\ :\ c_i(\sigma) \geq r\,\big\}\,, \qquad 3.2.42$$

consisting of the entries of σ that have at least r smaller labels to their right. The reason for this is best understood by working on an example. Note that for $\sigma = 72381645$ we have the following circle diagram



For convenience we have placed the entries of σ on top of their columns. From this it is easy to see that we have

$$M_{\geq 1} = \{2,3,6,7,8\},\ M_{\geq 2} = \{6,7,8\},\ M_{\geq 3} = \{7,8\},\ M_{\geq 4} = \{7,8\},\ M_{\geq 5} = \{7\},\ M_{\geq 6} = \{7\}$$

Note also that in this case

$$\lambda(\sigma) = (6,4,2,1,1) \quad , \quad \lambda'(\sigma) = (5,3,2,2,1,1) \quad \text{and} \quad l(\sigma) = 14\,.$$

Since in general

$$\lambda_i'(\sigma) \;=\; \sum_j \chi\big(\lambda_j(\sigma) \geq r\big)$$

we see that the successive sizes of the collections $M_{\geq r}(\sigma)$ give the components of $\lambda'(\sigma)$. Now in this case for the word $w = a_1 a_2 \cdots a_{14} \in RED(\sigma)$ to produce the monomial

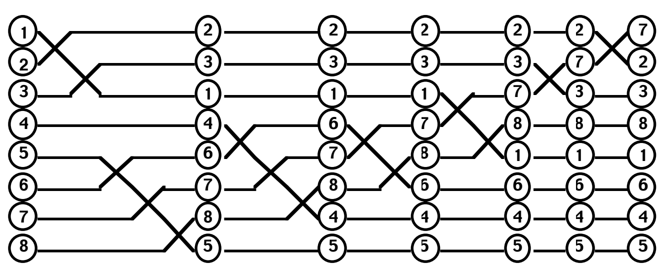
$$x_{\beta_1} x_{\beta_2} \cdots x_{\beta_{14}} \;=\; x_1^5\, x_2^3\, x_3^2\, x_4^2\, x_5^1\, x_6^1 \qquad 3.2.43$$

we must have

$$a_1 < a_2 < a_3 < a_4 < a_5\ ,\quad a_6 < a_7 < a_8\ ,\quad a_9 < a_{10}\ ,\quad a_{11} < a_{12}\ ,$$

Thus at the end of the first stage, 5 high labels will be involved. Each these labels will then have at least one smaller label to their right in the target permutation. But there are altogether only 5 such labels in our σ and they are precisely $2,3,6,7,8$. So the high labels involved in the first stage must be the elements of $M_{\geq 1}(\sigma)$. Similarly, the second stage must involve 3 high labels. Moreover, these labels must be a subset of the previous ones for otherwise there would be more than 5 entries of σ with at least one smaller element on their right. This means that each of the high labels involved in the second stage will have at least 2 smaller labels to their right in the target permutation. But σ has only 3 entries with this property and they are $6,7,8$. Thus again we see that the high labels involved in the second stage must be the elements of $M_{\geq 2}(\sigma)$. This reasonning forces the high labels involved in each stage to be a subset of the high labels involved in the previous stage. This forces the high labels involved in the r^{th} stage to be the elements of $M_{\geq r}(\sigma)$ precisely as asserted. It is easy to see that this argument, in full generality yields that there can be one and only one word $w \in RED(\sigma)$ yielding the monomial in 3.2.39 when $\mu = \lambda'(\sigma)$. This completes our proof.

For sake of completeness we include below the line diagram of the word that produces the monomial in 3.2.43 for the permuation 72681645



We terminate this section with a result that can be used to compute the Schur function expansion of the product of two or more Stanley symmetric functions.

Theorem 3.2.4

For $\alpha = \alpha_1\alpha_2\cdots\alpha_h \in S_h$ and $\beta = \beta_1\beta_2\cdots\beta_k \in S_k$ we have

$$F_\alpha \times F_\beta \quad = \quad F_{\alpha\otimes\beta} \tag{3.2.44}$$

with

$$\alpha\otimes\beta = \alpha_1\alpha_2\cdots\alpha_h(h+\beta_1)(h+\beta_2)\cdots(h+\beta_k) \tag{3.2.45}$$

Proof

Note that from3.2.45 we derive that we can obtain the reduced words of $\alpha\otimes\beta$ by taking pairs u, v with $u \in RED(\alpha)$ and $v \in RED(\beta)$ and then shuffling u with $h+v$. In symbols

$$RED\big(\alpha\otimes\beta)\big) \quad = \quad \bigcup_{u\in RED(\alpha)} \bigcup_{v\in RED(\beta)} u \sqcup\!\sqcup (h+v)\,.$$

Thus the definition in 3.2.1 gives

$$F_{\alpha\otimes\beta} \quad = \quad \sum_{u\in RED(\alpha)} \sum_{v\in RED(\beta)} \sum_{w\in u \sqcup\!\sqcup (h+v)} Q_{p(w)}\,. \tag{3.2.45}$$

The last summation should remind us of the expression occurring in 3.1.26. It develops that we can still use Theorem 3.2.4 here even though we are shuffling pairs of words rather than pairs of permutations. Briefly, the idea is to replace u and v by permutations of $1, 2, \ldots, l(\alpha)$ and $1, 2, \ldots, l(\beta)$ respectively by the standard procedure that preserves descents and then apply formula 3.1.26 to the resulting pair. In this manner we derive that

$$\sum_{w\in u \sqcup\!\sqcup (h+v)} Q_{p(w)} \quad = \quad Q_{p(u)} \times Q_{p(v)}$$

Substituting this into 3.2.45 gives

$$F_{\alpha\otimes\beta} \quad = \quad \sum_{u\in RED(\alpha)} \sum_{v\in RED(\beta)} Q_{p(u)} \times Q_{p(v)} \quad = \quad \Big(\sum_{u\in RED(\alpha)} Q_{p(u)}\Big) \times \Big(\sum_{v\in RED(\beta)} Q_{p(v)}\Big)$$

and thus 3.2.44 follows from the definition in 3.2.1.

Remark 3.2.2

Note that by taking α and β both Grassmanian we can use Theorem 3.2.4 in conjunction with Theorem 2.4.4 to obtain the Schur function expansion of the product of $S_{\lambda'(\alpha)}$ by $S_{\lambda'(\beta)}$. On the basis of this fact Stanley observed in [] that there could not be a rule simpler that that of Littlewood-Richardson to compute the Schur function expansion of an arbitrary F_σ. We believe however that the LR tree construction is conceptionally and algorithmically simpler (although not necessarily more efficient) than the LR-rule. What appears to have escaped from Stanley's reasoning is that the computation of product of Schur functions within the family of Stanley symmetric functions should in fact be easier since it may go through inductive steps involving a wider collection of functions. Indeed, the variety of possible circle diagrams is considerably wider than that of skew diagrams since all of the latter can be already be obtained from a circle diagrams of 321-avoiding permutations. What is also rather curious is that Lascoux and Schützenberger in [] herald their tree algorithm as an improvement (in efficiency) over the LR rule (which is quite untrue) and fail to notice that it is more elementary (see []) and that it applies to a wider class of symmetric functions, namely the Stanley symmetric functions.

3.3. Divided Differences and Schubert Polynomials.

We shall deal here with a family of divided difference operators δ_i (for $i = 1, 2, 3, \ldots$) acting on polynomials (or formal power series) in the variables $x_1, x_2, x_3, \ldots$. The definition of δ_i is quite simple. Namely we set

$$\delta_i \;=\; \delta_{x_i x_{i+1}}\,, \tag{3.3.1}$$

where for any polynomial P in the variables x, y

$$\delta_{xy} P(x,y) = \frac{P(x,y) - P(y,x)}{x-y}\,. \tag{3.3.2}$$

Note that we may also write this in the form

$$\delta_{xy} \;=\; \frac{1}{x-y}\big(1 - s_{xy}\big) \tag{3.3.3}$$

where s_{xy} is the operator that interchanges x and y. In particular we have

$$\delta_i \;=\; \frac{1}{x_i - x_{i+1}}\big(1 - s_i\big) \tag{3.3.4}$$

where $s_i = s_{x_i x_{i+1}}$ interchanges x_i and x_{i+1}.

Since δ_i acts only on the variables x_i, x_{i+1} to compute its action we only need to know the following identity

Proposition 3.3.1

$$\delta_i\, x_i^a x_{i+1}^b \;=\; \begin{cases} x_i^{a-1}x_{i+1}^b + \cdots + x_i^{a-r-1}x_{i+1}^{b+r} + \cdots + x_i^b x_{i+1}^{a-1} & \text{if } a > b \\ = 0 & \text{if } a = b \\ x_i^a x_{i+1}^{b-1} + \cdots + x_i^{a+r}x_{i+1}^{b-r-1} + \cdots + x_i^{b-1}x_{i+1}^a & \text{if } b > a \end{cases} \tag{3.3.5}$$

Proof

If $a > b$ we may write

$$\delta_i\, x_i^a x_{i+1}^b \;=\; \frac{x_i^a x_{i+1}^b - x_i^b x_{i+1}^a}{x_i - x_{i+1}} \;=\; (x_i x_{i+1})^b\, \frac{x_i^{a-b} - x_{i+1}^{a-b}}{x_i - x_{i+1}}$$

Thus

$$\delta_i\, x_i^a x_{i+1}^b \;=\; (x_i x_{i+1})^b \left(x_i^{a-b-1} + \cdots + x_i^{a-b-r-1}x_{i+1}^r + \cdots + x_{i+1}^{a-b-1}\right).$$

This proves the first identity in 3.3.5. The third identity follows in a similar way. The second one is trivial.

Proposition 3.3.2

These operators satisfy the following version of the "Leibnitz rule: "

$$a) \quad \delta_i\,(f\,g) \;=\; (\delta_i\,f)\,g \;+\; (s_i\,f)\,\delta_i\,g$$

In particular for f homogeneous of degree 1 we get

$$b) \quad \delta_{a_1}\delta_{a_2}\cdots\delta_{a_k}\,(f\,g) = \sum_{i=1}^{k}\big(\delta_{a_i}s_{a_{i+1}}\cdots s_{a_k}f\big)\;\delta_{a_1}\cdots[\delta_{a_i}]\cdots\delta_{a_k}g + \big(s_{a_1}s_{a_2}\cdots s_{a_k}f\big)\;\delta_{a_1}\delta_{a_2}\cdots\delta_{a_k}g \qquad 3.3.6$$

where $[\delta_{a_i}]$ indicates omission of the factor "δ_{a_i}".

Proof

From 3.3.5 we derive that

$$\delta_i\,(f\,g) \;=\; \frac{1}{x_i - x_{i+1}}\Big(((1-s_i)f)\,g \;+\; (s_i f)\,(1-s_i)\,g\Big),$$

and this is another way of writing 3.3.6. This proves 3.3.6 a) and the case $k=1$ of 3.3.6 b). Proceeding by induction on k assume 3.3.6 b) true for k. This given note that from 3.3.6 a) we get that

$$\begin{aligned}\delta_{a_1}\delta_{a_2}\cdots\delta_{a_{k+1}}\,(f\,g) &= \delta_{a_1}\delta_{a_2}\cdots\delta_{a_k}\big((\delta_{a_{k+1}}f)\,g + (s_{a_{k+1}}f)\,\delta_{a_{k+1}}g\big) = \\ &= (\delta_{a_{k+1}}f)\,\delta_{a_1}\delta_{a_2}\cdots\delta_{a_k}g + \sum_{i=1}^{k}\big(\delta_{a_i}s_{a_{i+1}}\cdots s_{a_k}(s_{a_{k+1}}f)\big)\,\delta_{a_1}\cdots[\delta_{a_i}]\cdots\delta_{a_k}\delta_{a_{k+1}}g \\ &\quad + \;\big(s_{a_1}s_{a_2}\cdots s_{a_k}(s_{a_{k+1}}f)\big)\;\delta_{a_1}\delta_{a_2}\cdots\delta_{a_k}\delta_{a_{k+1}}g\end{aligned}$$

This completes the induction and the proof of case b).

Most importantly we also have the so-called "*Nil Coxeter*" relations:

Proposition 3.3.3

$$\begin{aligned} &i) \quad \delta_i\,\delta_i = 0 \qquad (\forall\;\; i \geq 1\,) \\ &ii) \quad \delta_i\;\delta_{i+1}\delta_i = \delta_{i+1}\,\delta_i\;\delta_{i+1} \\ &iii) \quad \delta_i\,\delta_j = \delta_j\,\delta_i \qquad (\forall\;\;|i-j| \geq 2)\end{aligned} \qquad 3.3.7$$

Proof

It follows immediately from the definition in 3.3.4 that δ_i kills every symmetric function of x_i, x_{i+1}. Thus, since in each of the three cases in 3.3.5 the result is symmetric, we derive that

$$\delta_i^2\,x_i^a x_{i+1}^b \;=\; 0\,.$$

This proves 3.3.7 i). The identity in 3.3.7 iii) is trivial since when $|i-j| \geq 2$ the two operators δ_i and δ_j act on disjoint sets of indices. The identity in 3.3.5 ii) is proved by noticing that repeated uses of 3.3.4 give

$$\delta_1\delta_2\delta_1 \;=\; \delta_2\delta_1\delta_2 \;=\; \frac{1}{(x_1-x_2)(x_1-x_3)(x_2-x_3)}\sum_{\sigma\in S_3} sign(\sigma)\sigma\,. \qquad 3.3.8$$

It follows from 3.3.7 $i)$ that for $w = a_1 a_2 \cdots a_l$ we shall have

$$\delta_{a_1}\delta_{a_2}\cdots\delta_{a_l} \neq 0$$

if and only if w is a reduced word of some permutation σ. Moreover, using $ii)$ and $iii)$ of 3.3.7 we can show that if $w = a_1 a_2 \cdots a_l$ and $w' = a'_1 a'_2 \cdots a'_l$ are both reduced words for the same permutation σ then we necessarily have

$$\delta_{a_1}\delta_{a_2}\cdots\delta_{a_l} \;=\; \delta_{a'_1}\delta_{a'_2}\cdots\delta_{a'_l}\ .$$

This means that to any $\sigma \in S_n$ we can associate a well defined divided difference operator δ_σ simply by setting for any reduced word $w = a_1 a_2 \cdots a_l \in RED(\sigma)$

$$\delta_\sigma \;=\; \delta_{a_1}\delta_{a_2}\cdots\delta_{a_l}\ .$$

Here and after the symbol $\sigma^{(n)}$ will denote the top permutation of s_n. That is

$$\sigma^{(n)} \;=\; \begin{bmatrix} 1 & 2 & \cdots & n \\ n & n-1 & \cdots & 1 \end{bmatrix}. \qquad 3.3.9$$

Remarkably, the operator corresponding to the top element is a version of complete "*symmetrization*". More precisely we have the following general form of 3.3.8.

Proposition 3.3.4

$$\delta_{\sigma^{(n)}} \;=\; \frac{1}{\prod_{1\le i<j\le n}(x_i - x_j)} \sum_{\sigma\in S_n} sign(\sigma)\,\sigma \qquad 3.3.10$$

Proof

The canonical factorization of $\sigma^{(n)}$ and 3.3.4 gives

$$\delta_{\sigma^{(n)}} = \prod_{i=1}^{n-1} \delta_{n-1}\delta_{n-2}\cdots\delta_i = \prod_{i=1}^{n-1} \frac{1}{x_{n-1}-x_n}(1-s_{n-1})\frac{1}{x_{n-2}-x_{n-1}}(1-s_{n-2})\cdots\frac{1}{x_i - x_{i+1}}(1-s_i) \qquad 3.3.11$$

where the factors are to be taken from left to right. This given we see that $\delta_{\sigma^{(n)}}$ is of the form

$$\delta_{\sigma^{(n)}} \;=\; \sum_{\sigma\in S_n} a_\sigma(x)\,\sigma \qquad 3.3.12$$

with the coefficients $a_\sigma(x)$ rational functions of $x_1, x_2, \ldots, x_n$. Now note that since the product

$$\delta_j \prod_{i=1}^{n-1} \delta_{n-1}\delta_{n-2}\cdots\delta_i$$

has $\binom{n}{2}+1$ factors, it does not correspond to any reduced factorization. Consequently we must have

$$\delta_j\,\delta_{\sigma^{(n)}} \;=\; 0. \qquad (\text{for } j = 1, 2, \ldots, n-1\ ).$$

In view of 3.3.4 this may also be written as

$$\delta_{\sigma^{(n)}} \;=\; s_j\,\delta_{\sigma^{(n)}} \qquad (\text{for } j = 1, 2, \ldots, n-1\ ).$$

It thus follows that we must also have

$$\delta_{\sigma^{(n)}} \;=\; \alpha\ \delta_{\sigma^{(n)}} \qquad (\ \forall \alpha \in S_n\ )$$

Using 3.3.12 this becomes

$$\sum_{\sigma \in S_n} (\alpha\, a_\sigma(x))\ \alpha\,\sigma \;=\; \sum_{\sigma \in S_n} a_\sigma(x)\,\sigma\,.$$

Equating coefficients of $\alpha\,\beta$ we get

$$\alpha\, a_\beta(x) \;=\; a_{\alpha\beta}(x) \qquad 3.3.13$$

This means that we only need to compute one of these coefficients. Now we see from 3.3.11 that

$$\begin{aligned} a_{\sigma^{(n)}}(x)\,\sigma^{(n)} &= \prod_{i=1}^{n-1} \frac{1}{x_{n-1}-x_n}(-s_{n-1})\frac{1}{x_{n-2}-x_{n-1}}(-s_{n-2})\cdots\frac{1}{x_i-x_{i+1}}(-s_i) \\ &= (-1)^{\binom{n}{2}}\Big(\prod_{i=1}^{n-1}\frac{1}{x_{n-i}-x_{n-i+1}}\cdots\frac{1}{x_1-x_{n-i+1}}\Big)\prod_{i=1}^{n-1} s_{n-1}s_{n-2}\cdots s_i \\ &= \frac{(-1)^{\binom{n}{2}}}{\prod_{1\le i<j\le n}(x_i-x_j)}\ \sigma^{(n)} \end{aligned}$$

So from 3.3.13 for $\alpha\beta = \sigma$ and $\beta = \sigma^{(n)}$ we get

$$a_\sigma(x) \;=\; \sigma\sigma^{(n)}\Big(\frac{(-1)^{\binom{n}{2}}}{\prod_{1\le i<j\le n}(x_i-x_j)}\Big) \;=\; \frac{sign(\sigma)}{\prod_{1\le i<j\le n}(x_i-x_j)}\,.$$

This proves 3.3.10.

This proposition has the following immediate corollary

Theorem 3.3.1

For any partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ we have

$$\delta_{\sigma^{(n)}}\big(x_1^{\lambda_1+n-1}x_2^{\lambda_2+n-2}\cdots x_1^{\lambda_n+n-n}\big) \;=\; S_\lambda(x_1, x_2, \ldots, x_n)\,. \qquad 3.3.14$$

The fact that a Schur function can be obtained by the action of the difference operator $\delta_{\sigma^{(n)}}$ on a monomial should suggest that an interesting family of polynomials might be obtained by the action of the general operators δ_σ. This is precisely the discovery of Lascoux and Schützenberger in []. In fact for a $\sigma \in S_n$ the Schubert polynomial $\mathcal{SC}_\sigma(x)$ is defined by setting

$$\mathcal{SC}_\sigma(x) \;=\; \delta_{\sigma^{-1}\sigma^{(n)}}\, x_1^{n-1}x_2^{n-2}\cdots x_{n-1}^{1}\,. \qquad 3.3.15$$

In particular we get

$$\mathcal{SC}_{\sigma^{(n)}}(x) \;=\; x_1^{n-1}x_2^{n-2}\cdots x_{n-1}^{1}\,. \qquad 3.3.16$$

The polynomials $\mathcal{SC}_\sigma(x)$ have been shown to have remarkable properties The reader will find a detailed presentation of basic results of theory of Schubert polynomials in Macdonald's book []. For sake of completeness we shall reproduce here the statements and proofs of the results that we will need in our proof of Theorem 1.4.1.

Remark 3.3.1

Note that for any $\sigma \in S_n$ we have

$$l\big(\sigma^{-1}\sigma^{(n)}\big) \;=\; \binom{n}{2} \;-\; l(\sigma) \qquad 3.17$$

the reason for this is that all the inversions of σ^{-1} are transformed into non-inversions after right multiplication by $\sigma^{(n)}$. Thus $l\big(\sigma^{-1}\sigma^{(n)}\big) = \binom{n}{2} - l(\sigma^{-1})$, and then 3.3.17 follows since $l(\sigma) = l(\sigma^{-1})$.

Note next that we can always find a sequence of indices $a_1a_2\cdots a_k$ with $1 \leq a_i \leq n-1$ such that

$$l(\sigma s_{a_1}s_{a_2}\cdots s_{a_i}) \;=\; l(\sigma)+i \qquad \text{for } i = 1,2,\ldots,k \qquad 3.18$$

and

$$\sigma s_{a_1}s_{a_2}\cdots s_{a_k} \;=\; \sigma^{(n)} \qquad 3.19$$

To do this we simply choose s_{a_i} to be any of the transpositions that interchanges two adjacent elements of $\sigma s_{a_1}s_{a_2}\cdots s_{a_{i-1}}$ that are in the right order. This will eventually bring us to the top element of S_n at which time we stop. Now 3.17, 3.18 for $i = k$ and 3.19 give

$$a)\quad k \;=\; \binom{n}{2} - l(\sigma) \qquad \text{and} \qquad b)\quad \sigma^{-1}\sigma^{(n)} = s_{a_1}s_{a_2}\cdots s_{a_k} \qquad 3.20$$

In particular from 3.20 a) we derive that

$$a_1a_2\cdots a_k \in RED(\sigma^{-1}\sigma^{(n)})\,.$$

This given, the definition in 3.3.15 yields

$$\mathcal{SC}_\sigma(x) \;=\; \delta_{a_1}\delta_{a_2}\cdots\delta_{a_k}\, x_1^{n-1}x_2^{n-2}\cdots x_{n-1}^{1}\,. \qquad 3.3.21$$

These observations immediately yield us the following two basic facts.

Theorem 3.3.2

For $\sigma \in S_n$, $\mathcal{SC}_\sigma(x)$ is a homogeneous polynomial of degree $l(\sigma)$ in $x_1, x_2, \ldots, x_{n-1}$.

Proof

We see from 3.3.5 that each δ_i preserves homogeneity and lowers degrees by 1. This given the statement follows from 3.20 a) and formula 3.3.21.

We shall here and after denote by $\mathcal{A}_n$ the collection of monomials

$$\mathcal{A}_n \;=\; \big\{x_1^{\epsilon_1}x_1^{\epsilon_1}\cdots x_{n-1}^{\epsilon_{n-1}} \,:\, 0 \leq \epsilon_i \leq n-i \quad \text{for} \quad i = 1,2,\ldots,n-1\big\}\,.$$

It is well known that $\mathcal{A}_n$ is a basis for the quotient

$$\mathbb{Q}[x_1, x_2, \ldots, x_n]/(e_1, e_2, \ldots, e_n)$$

where $e_1, e_2, \ldots, e_n$ are the elementary symmetric functions. It develops that Schubert polynomials may be integrally expanded in terms of these monomials. More precisely

Theorem 3.3.3

For $\sigma \in S_n$

$$\mathcal{SC}_\sigma(x) \;=\; \sum_{x^p \in \mathcal{A}_n} a_p\, x^p \qquad 3.3.22$$

where the coefficients a_p are non-negative integers. For the identity permutation this reduces to

$$\mathcal{SC}_I(x) \;=\; 1 \qquad 3.3.23$$

Proof

In view of formula 3.3.21, to prove 3.3.24 we need only show that each δ_i sends any element of $\mathcal{A}_n$ into a $\mathbb{N}$-linear combination of elements of $\mathcal{A}_n$. However this follows immediately from formula 3.3.5. In fact, if $a = \epsilon_i$ and $b = \epsilon_{i+1}$ we get that

$$\delta_i\; x_i^{\epsilon_i} x_{i+1}^{\epsilon_{i+1}}$$

is a sum of monomials of the form

$$x_i^{\epsilon_i - r - 1} x_{i+1}^{\epsilon_{i+1}+r} \qquad (\text{with } r \geq 0 \quad \text{and} \quad \epsilon_{i+1} + r \leq \epsilon_i - 1)$$

if $\epsilon_i > \epsilon_{i+1}$ or a sum of monomials of the form

$$x_i^{\epsilon_i + r} x_{i+1}^{\epsilon_{i+1} - r - 1} \qquad (\text{with } r \geq 0 \quad \text{and} \quad \epsilon_i + r \leq \epsilon_{i+1} - 1)$$

if $\epsilon_i < \epsilon_{i+1}$. In either case we see that $\epsilon_i \leq n - i$ and $\epsilon_{i+1} \leq n - i - 1$ force all these summands to be of the form

$$x_i^{p_i} x_{i+1}^{p_{i+1}} \qquad (\text{with } \; p_i \leq n - i \quad \text{and} \quad p_{i+1} \leq n - i - 1$$

and this is all that is needed to show the first assertion of the Theorem. To complete the proof we note that by definition we have

$$\mathcal{SC}_I(x) \;=\; \delta_{\sigma^{(n)}} x_1^{n-1} x_2^{n-1} \cdots x_{n-1}^1 \,,$$

but then 3.3.10 gives

$$\mathcal{SC}_I(x) = \frac{1}{\prod_{1 \leq i < j \leq n}(x_i - x_j)} \sum_{\sigma \in S_n} sign(\sigma)\, x_{\sigma_1}^{n-1} x_{\sigma_2}^{n-1} \cdots x_{\sigma_{n-1}}^1 = \frac{\prod_{1 \leq i < j \leq n}(x_i - x_j)}{\prod_{1 \leq i < j \leq n}(x_i - x_j)} = 1$$

This proves the second assertion.

The following identities enable us to obtain explicit expressions for some Schubert polynomials.

Proposition 3.3.4

For $u, \sigma \in S_n$

$$\delta_u \, \mathcal{SC}_\sigma(x) \;=\; \begin{cases} \mathcal{SC}_{\sigma u^{-1}}(x) & \textit{if } \; l(\sigma u^{-1}) = l(\sigma) - l(u) \\ 0 & \textit{otherwise} \end{cases} \tag{3.3.24}$$

In particular when $1 \leq i \leq n-1$

$$\delta_i \; \mathcal{SC}_\sigma(x) \;=\; \begin{cases} \mathcal{SC}_{\sigma s_i}(x) & \text{if } \sigma_i > \sigma_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{3.3.25}$$

Proof

From the definition we get

$$\delta_u \, \mathcal{SC}_\sigma(x) \;=\; \delta_u \, \delta_{\sigma^{-1}\sigma^{(n)}} \, x_1^{n-1} x_2^{n-2} \cdots x_2$$

Now clearly $\delta_u \, \delta_{\sigma^{-1}\sigma^{(n)}} = 0$ unless

$$l(u) \;+\; l\big(\sigma^{-1}\sigma^{(n)}\big) \;=\; l\big(u\,\sigma^{-1}\sigma^{(n)}\big) \;=\; l\big((\sigma\, u^{-1})^{-1}\sigma^{(n)}\big) \tag{3.3.26}$$

in which case

$$\delta_u \, \delta_{\sigma^{-1}\sigma^{(n)}} \;=\; \delta_{(\sigma\, u^{-1})^{-1}\sigma^{(n)}} \, .$$

However, from 3.17 we derive that 3.3.26 is equivalent to

$$l(u) \;+\; \binom{n}{2} \;-\; l(\sigma) \;=\; \binom{n}{2} \;-\; l\big(\sigma\, u^{-1}\big) \, ,$$

or better

$$l\big(\sigma\, u^{-1}\big) \;=\; l(\sigma) \;-\; l(u) \, .$$

This proves 3.24. In particular we get

$$\delta_i \; \mathcal{SC}_\sigma(x) \;=\; \begin{cases} \mathcal{SC}_{\sigma s_i}(x) & \text{if } l(\sigma s_i) = l(\sigma) - 1 \\ 0 & \text{otherwise} \end{cases}$$

and 3.3.25 then follows since $l(\sigma s_i) = l(\sigma) - 1$ holds if and only if $\sigma_i > \sigma_{i+1}$. This completes the proof.

Schubert polynomials have several interesting properties the following two are worth mentioning here

Theorem 3.3.4

For any $\sigma \in S_n$

a) $\mathcal{SC}_\sigma(x)$ *is symmetric in* $x_i \, , x_{i+1}$ *if and only if* $\sigma_i < \sigma_{i+1}$

b) If $1 \leq r < n$ *is the last descent of* σ *then* $\mathcal{SC}_\sigma(x) \in \mathbb{N}[x_1, x_2, \ldots, x_r]$

Proof

Formula 3.3.24 yields that

$$\delta_i \; \mathcal{SC}_\sigma(x) \;=\; 0 \tag{3.3.27}$$

if and only if $\sigma_i < \sigma_{i+1}$. However 3.3.4 shows that 3.3.27 is equivalent to

$$\mathcal{SC}_\sigma(x) \;=\; s_i\,\mathcal{SC}_\sigma(x)\,.$$

This proves the assertion in a). Note next that if r is the last descent, then

$$\sigma_{r+1} < \sigma_{r+2} < \cdots < \sigma_n$$

So part a) gives that $\mathcal{SC}_\sigma(x)$ is symmetric in $x_{r+1}, x_{r+2}, \ldots, x_n$. But from Theorem 3.3.2 it follows that $\mathcal{SC}_\sigma(x)$ does not depend on x_n, Therefore it cannot depend on $x_{r+1}, x_{r+2}, \ldots, x_{n-1}$ as well. This proves part b).

Let $\mathcal{H}_n$ denote the linear span of the monomials in $\mathcal{A}_n$, in symbols

$$\mathcal{H}_n \;=\; \mathcal{L}\big[x_1^{\epsilon_1}x_2^{\epsilon_1}\cdots x_{n-1}^{\epsilon_{n-1}} \;:\; 0 \le \epsilon_i \le n-i\big]\,. \qquad 3.3.28$$

This given we have the following useful result.

Theorem 3.3.5

The collection $\{\mathcal{SC}_\sigma(x)\}_{\sigma\in S_n}$ *is a basis of* $\mathcal{H}_n$ *and for any polynomial* $P \in \mathcal{H}_n$ *we have the expansion formula*

$$P(x_1, x_2, \ldots, x_{n-1}) \;=\; \sum_{\sigma\in S_n} \delta_\sigma P\,\big|_{x=0}\, \mathcal{SC}_\sigma(x_1, x_2, \ldots, x_{n-1}) \qquad 3.3.29$$

Proof

The definition in 3.28 gives that

$$dim\,\mathcal{H}_n \;=\; n! \;=\; \#\big\{\mathcal{SC}_\sigma(x)\big\}_{\sigma\in S_n}\;.$$

Since Theorem 3.3.2 gives $\big\{\mathcal{SC}_\sigma(x)\big\}_{\sigma\in S_n} \subseteq \mathcal{H}_n$, we need only show independence. To this end let

$$P(x) \;=\; \sum_{\sigma\in S_n} a_\sigma\,\mathcal{SC}_\sigma(x) \qquad 3.3.30$$

Note that the homogeneity of $\mathcal{SC}_\sigma(x)$ coupled with formulas 3.3.23 and 3.3.24 give

$$\delta_\alpha\,\mathcal{SC}_\sigma(x)\,\big|_{x=0} \;=\; \begin{cases} 1 & \text{if } \alpha=\sigma, \\ 0 & \text{otherwise.} \end{cases} \qquad 3.3.31$$

Applying δ_α to 3.3.30 and setting $x = 0$ we get

$$a_\alpha \;=\; \delta_\alpha\,P\,\big|_{x=0}\,. \qquad 3.3.32$$

Thus $P = 0 \;\Rightarrow\; a_\alpha = 0$, proving independence. This given, 3.29 follows from 3.3.32.

The following beautiful result of Billey, Jockusch and Stanley reveals the intimate relationship between Schubert polynomials and Stanley symmetric functions.

Theorem 3.3.5

For any permutation $\sigma \in S_n$ of length l we have

$$\mathcal{SC}_\sigma(x) = \sum_{a_1 a_2 \cdots a_l \in RED(\sigma)} \sum_{\substack{1 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_l \leq n-1 \\ a_i < a_{i+1} \Rightarrow \beta_i < \beta_{i+1} \\ \beta_i \leq a_i}} x_{\beta_1} x_{\beta_2} \cdots x_{\beta_l} \qquad 3.3.33$$

We shall give here the remarkably simple proof of this result due to Fomin and Stanley []. To this end we need to present some auxiliary material. To begin we note that the right hand side of this identity has a very simple expression in terms of the Nil-Coxeter algebra.

Proposition 3.3.5

For any $\sigma \in S_n$ we have

$$A_1(x_1)A_2(x_2)\cdots A_{n-1}(x_{n-1})\Big|_{u_\sigma} = \sum_{a_1 a_2 \cdots a_l \in RED(\sigma)} \sum_{\substack{1 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_l \leq n-1 \\ a_i < a_{i+1} \Rightarrow \beta_i < \beta_{i+1} \\ \beta_i \leq a_i}} x_{\beta_1} x_{\beta_2} \cdots x_{\beta_l} \qquad 3.3.34$$

Proof

It is easily seen from the definition in 3.2.14 that the expansion of the product on the left hand side produces terms of the form

$$x_{\beta_1} x_{\beta_2} \cdots x_{\beta_m}\, u_{a_1} u_{a_2} \cdots u_{a_m}\Big|_{u_\sigma}$$

with

$$\beta_1 \leq \beta_2 \leq \cdots \leq \beta_m$$

satisfying

$$a_i < a_{i+1} \implies \beta_i < \beta_{i+1}$$

this is for the same reason as in the proof of 3.2.22. However in this case we have the additional feature that the factor $A_\beta(x_\beta)$ contributes only terms $x_\beta\, u_a$ with $a \geq \beta$. This shows that we must also have the inequalities

$$\beta_i \leq a_i \qquad (\text{ for } i = 1, 2, \ldots, m\ ).$$

Now the Nil-Coxeter relations in 3.2.12 again guarantee that the only terms that survive are those for which $m = l$ and

$$u_{a_1} u_{a_2} \cdots u_{a_l} \;=\; u_\sigma\,.$$

This completes the proof of 3.3.34.

To proceed we need one more identity of the Nil-Coxeter algebra.

Proposition 3.3.6

For any $1 \le i < n$ we have

$$\delta_{xy}\,\mathcal{A}_i(x)\mathcal{A}_{i+1}(y) \;=\; \mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i \qquad 3.3.35$$

Proof

Note that from the definition in 3.2.14 we get that

$$\mathcal{A}_i(x)\mathcal{A}_i(y) \;=\; \mathcal{A}_i(x)\mathcal{A}_{i+1}(y) \;+\; y\,\mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i \qquad 3.3.36$$

Interchanging x and y gives

$$\mathcal{A}_i(y)\mathcal{A}_i(x) \;=\; \mathcal{A}_i(y)\mathcal{A}_{i+1}(x) \;+\; x\,\mathcal{A}_i(y)\mathcal{A}_{i+1}(x)\,u_i \qquad 3.3.37$$

Subtracting 3.3.37 from 3.3.36 and using Proposition 3.2.1 we get

$$\delta_{xy}\,\mathcal{A}_i(x)\mathcal{A}_{i+1}(y) \;=\; \frac{\big(\,x\mathcal{A}_i(y)\mathcal{A}_{i+1}(x)\,u_i \;-\; y\,\mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i\,\big)}{x-y} \qquad 3.3.38$$

But we have

$$\begin{aligned}
\mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i &= \mathcal{A}_{i+1}(x)(1+xu_i)\mathcal{A}_{i+1}(y)\,u_i \\
&= \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y) \,+\, x\,\mathcal{A}_{i+1}(x)\,u_i\,\mathcal{A}_{i+1}(y)\,u_i \\
&= \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y) \,+\, x\,\mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)\,u_i\,(1+y\,u_{i+1})\,u_i \\
&= \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y) \,+\, xy\,\mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)\,u_i\,u_{i+1}\,u_i \\
&= \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y) \,+\, xy\,\mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)\,u_{i+1}\,u_i\,u_{i+1} \\
&= \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y) \,+\, xy\,\mathcal{A}_{i+1}(x)\mathcal{A}_{i+2}(y)(1+y\,u_{i+1})\,u_{i+1}\,u_i\,u_{i+1} \\
&= \mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y) \,+\, xy\,\mathcal{A}_{i+1}(x)\mathcal{A}_{i+1}(y)\,u_{i+1}\,u_i\,u_{i+1}
\end{aligned}$$

Since this last expression is completely symmetric in x and y (again by Proposition 3.2.1) we deduce that

$$\mathcal{A}_i(y)\mathcal{A}_{i+1}(x)\,u_i \;=\; \mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i\,.$$

Using this in 3.3.38 gives that

$$\delta_{xy}\,\mathcal{A}_i(x)\mathcal{A}_{i+1}(y) \;=\; \frac{\big(\,x\mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i \;-\; y\,\mathcal{A}_i(x)\mathcal{A}_{i+1}(y)\,u_i\,\big)}{x-y}$$

which is easily seen to simplify to 3.3.35.

We now have all the ingredients we need to establish the Billey-Jockusch-Stanley formula.

Proof of Theorem 3.3.5

For convenience, let us for a moment denote by $G_\sigma(x)$ the right hand side of 3.3.33. Recalling in extent the definition in 3.2.14, 3.3.34 gives

$$\begin{aligned} G_\sigma(x) = (1+x_1u_{n-1})(1+x_1u_{n-2})\cdots(1+x_1u_1)\times \\ (1+x_2u_{n-1})(1+x_2u_{n-2})\cdots(1+x_2u_2)\times \\ \cdots \qquad \cdots \qquad \cdots \\ (1+x_{n-2}u_{n-1})(1+x_{n-2}u_{n-2})\times \\ (1+x_{n-1}u_{n-1})\Big|_{u_\sigma} \end{aligned} \tag{3.3.39}$$

A view at this display makes it palpably clear that the only way to obtain a term involving $u_{\sigma^{(n)}}$ from this expression is to pick the "x" part in every one of the factors. Thus we must have

$$G_{\sigma^{(n)}}(x) \;=\; x_1^{n-1}x_2^{n-2}\cdots x_{n-1}^1\,.$$

This proves 3.3.33 for the top permutation. We can thus proceed by descent induction on the length of σ. Let us then assume that we have proved $G_\sigma(x)=\mathcal{SC}_\sigma(x)$ for all $\sigma\in S_n$ of length $l+1$ and let $\alpha\in S_n$ be of length l. Since α is not the top element there will be an index $i<n$ for which $\alpha_i<\alpha_{i+1}$. This gives that the permutation αs_i has length $l+1$ so by the induction hypothesis we have

$$G_{\alpha s_i}(x) \;=\; \mathcal{SC}_{\alpha s_i}(x)\,.$$

Now 3.3.25 can be applied to $\sigma=\alpha s_i$ and obtain

$$\mathcal{SC}_\alpha(x) \;=\; \delta_i\,\mathcal{SC}_{\alpha s_i}(x) \;=\; \delta_i\,G_{\alpha s_i}(x)\,.$$

Now, using 3.3.39 we get that

$$\begin{aligned} \mathcal{SC}_\alpha(x) \;&=\; \delta_i\mathcal{A}_1(x_1)\cdots\mathcal{A}_i(x_i)\mathcal{A}_{i+1}(x_{i+1})\cdots\mathcal{A}_{n-1}(x_{n-1})\Big|_{u_\alpha u_i} \\ &=\; \mathcal{A}_1(x_1)\cdots(\delta_i\mathcal{A}_i(x_i)\mathcal{A}_{i+1}(x_{i+1}))\cdots\mathcal{A}_{n-1}(x_{n-1})\Big|_{u_\alpha u_i} \\ \text{(using 3.3.35)} \;&=\; \mathcal{A}_1(x_1)\cdots(\mathcal{A}_i(x_i)\mathcal{A}_{i+1}(x_{i+1})u_i)\cdots\mathcal{A}_{n-1}(x_{n-1})\Big|_{u_\alpha u_i} \\ \text{(using 3.2.12 b))} \;&=\; \mathcal{A}_1(x_1)\cdots\mathcal{A}_i(x_i)\mathcal{A}_{i+1}(x_{i+1})\cdots\mathcal{A}_{n-1}(x_{n-1})\;u_i\Big|_{u_\alpha u_i} \\ &=\; \mathcal{A}_1(x_1)\mathcal{A}_2(x_2)\cdots\mathcal{A}_{n-1}(x_{n-1})\Big|_{u_\alpha} \;=\; G_\alpha(x)\,. \end{aligned}$$

This completes the induction and the proof of the Theorem.

An immediate corollary of Theorem 3.3.5 is the following important identity.

Theorem 3.3.6

If $\sigma\in S_n$ is any permutation of length l then

$$\mathcal{SC}_{1_m\otimes\sigma}(x_1,x_2,\ldots,x_n) \;=\; F_\sigma(x_1,x_2,\ldots,x_n) \qquad (\forall \quad m\geq n) \tag{3.3.40}$$

Proof

Note that we have

$$a_1a_2\cdots a_l \in RED(\sigma) \quad \longleftrightarrow \quad a_1+m\ a_2+m\ \cdots\ a_l+m\ \in RED(1_m\otimes\ \sigma)\,.$$

Thus formula 3.3.33 for $1_m\otimes\ \sigma$ may be written in the form

$$\mathcal{SC}_{1_m\otimes\,\sigma}\,(x_1,x_2,\ldots,x_{m+n-1}) \;=\; \sum_{a_1a_2\cdots a_l\in RED(\sigma)}\;\; \sum_{\substack{1\le\beta_1\le\beta_2\le\cdots\le\beta_l\le m+n-1\\ a_i+m<a_{i+1}+m\;\Rightarrow\;\beta_i<\beta_{i+1}\\ \beta_i\le a_i+m}} x_{\beta_1}x_{\beta_2}\cdots x_{\beta_l}\,. \qquad 3.3.41$$

But if all the variables $x_{n+1},x_{n+2},\ldots,x_{m+n-1}$ are set to zero the condition $\beta_i \le \alpha_i+m$ becomes vacuous when $m\ge n$, and so 3.3.41 yields

$$\mathcal{SC}_{1_m\otimes\,\sigma}\,(x_1,x_2,\ldots,x_{m+n-1})\Big|_{x_{n+1},\ldots,x_{m+n-1}=0} = \sum_{a_1a_2\cdots a_l\in RED(\sigma)}\;\; \sum_{\substack{1\le\beta_1\le\beta_2\le\cdots\le\beta_l\le n\\ a_i<a_{i+1}\;\Rightarrow\;\beta_i<\beta_{i+1}}} x_{\beta_1}x_{\beta_2}\cdots x_{\beta_l}\,.$$

This proves 3.3.40.

Before we proceed any further it will be good to note that Schubert polynomials are stable under the natural embedding of S_n into S_{n+m}. To be precise we have the following general result.

Proposition 3.3.7

If $\sigma=\sigma_1\sigma_2\cdots\sigma_n\in S_n$ has last descent at r then for any $m\ge 0$ we have

$$\mathcal{SC}_{\sigma\otimes 1_m}(x_1,x_2,\ldots,x_r) \;=\; \mathcal{SC}_{\sigma}(x_1,x_2,\ldots,x_r)\,. \qquad 3.3.42$$

Proof

By definition

$$\sigma\otimes 1_m \;=\; \begin{bmatrix} 1 & 2 & \cdots & n & n+1 & n+2 & \cdots & n+m\\ \sigma_1 & \sigma_2 & \cdots & \sigma_n & n+1 & n+2 & \cdots & n+m\end{bmatrix},$$

In particular also $\sigma\otimes 1_m$ has last descent at r. Thus from Theorem 3.3.4 we derive that both sides of 3.3.42 are polynomials in $x_1,x_2,\ldots,x_r$. Moreover we see that we also trivially have

$$RED(\sigma) \;=\; RED(\sigma\otimes 1_m)\,.$$

Thus 3.3.42 follows immediately from Theorem 3.3.5.

This given, here and after we will make replacements $\sigma\ \to\sigma\otimes 1_m$, whenever necessary to keep all the permutations, indexing Schubert polynomials appearing in a given identity, in the same Symmetric Group. Keeping this in mind we have the following basic result.

Theorem 3.3.7

For any $u\in S_n$ we have

$$(\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n)\mathcal{SC}_u(x_1,x_2,\ldots,x_{n-1}) = \sum_{\substack{1\le a<b\le n+1\\ l(u\times t_{ab})=l(u)+1}} (\alpha_a-\alpha_b)\ \mathcal{SC}_{u\times t_{ab}}(x_1,x_2,\ldots,x_n) \qquad 3.3.43$$

Proof

Since by Theorem 3.3.3 we have $\mathcal{SC}_u(x_1, x_2, \ldots, x_{n-1}) \in \mathcal{H}_n$, it follows that the left hand side of 3.3.43 is in $\mathcal{H}_{n+1}$. We can thus apply Theorem 3.3.35 and obtain the expansion

$$f\,\mathcal{SC}_u(x_1, x_2, \ldots, x_{n-1}) = \sum_{\sigma \in S_{n+1}} \delta_\sigma(f\mathcal{SC}_u)\big|_{x=0} \mathcal{SC}_\sigma(x_1, x_2, \ldots, x_n) \qquad 3.3.44$$

where for convenience we have set

$$f = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n\,. \qquad 3.3.45$$

Assuming that $l(u) = l - 1$, it follows that the product $f\mathcal{SC}_u$ is a homogeneous polynomial of degree l and therefore the summation in 3.3.44 need only be carried out over permutations σ of length l. This given assuming that

$$a_1 a_2 \cdots a_l \in RED(\sigma)$$

we may compute the coefficient of $\mathcal{SC}_\sigma$ in 3.3.34 by means of formula 3.3.6 with $g = \mathcal{SC}_u$ and $k = l$. We thus obtain

$$\delta_\sigma(f\mathcal{SC}_u)\big|_{x=0} \;=\; \sum_{i=1}^{l} \big(\delta_{a_i s_{a_{i+1}} \cdots s_{a_l}} f\big)\; \delta_{a_1} \cdots [\delta_{a_i}] \cdots \delta_{a_l} \mathcal{SC}_u\,. \qquad 3.3.46$$

Note that we need not evaluate at $x = 0$ on the right hand side here since $\mathcal{SC}_u$ is homogeneous of degree $l - 1$. For the same reason we have

$$\delta_{a_1}\delta_{a_2} \cdots \delta_{a_l}\, \mathcal{SC}_u \;=\; 0\,,$$

so no additional term is needed in 3.3.46. Now it follows from formula 3.3.24 that we have

$$\delta_{a_1} \cdots [\delta_{a_i}] \cdots \delta_{a_l}\, \mathcal{SC}_u \;=\; \begin{cases} 1 & \text{if } a_1 \cdots [a_i] \cdots a_l \in RED(u)\,, \\ 0 & \text{otherwise}\,. \end{cases} \qquad 3.3.47$$

Now it is easy to see that if

$$(a, b) \;=\; s_{a_l} \cdots s_{a_{i+1}}\,(a_i\,, a_i + 1) \qquad 3.3.48$$

then

$$t_{ab} \;=\; s_{a_l} \cdots s_{a_i} \cdots s_{a_l} \qquad \text{and} \qquad s_{a_1} \cdots [s_{a_i}] \cdots s_{a_l} \;=\; \sigma \times t_{ab} \qquad 3.3.49$$

and thus from 3.3.47 we deduce that the only terms that survive in 3.3.44 are those for which

$$\sigma = u \times t_{ab} \qquad \text{and} \qquad l(u \times t_{ab}) = 1\,.$$

for some $1 \le a < b \le n + 1$. This given note that 3.3.48 gives that

$$s_{a_l} \cdots s_{a_{i+1}} \delta_{a_i} s_{a_{i+1}} \cdots s_{a_l} \;=\; \delta_{x_a, x_b}$$

Thus from 3.3.45 we get that

$$\delta_{a_i s_{a_{i+1}} \cdots s_{a_l}} f \;=\; s_{a_l} \cdots s_{a_{i+1}} \delta_{a_i} s_{a_{i+1}} \cdots s_{a_l} f \;=\; \delta_{x_a, x_b} f \;=\; \alpha_a - \alpha_b\,.$$

In summary when we have 3.3.49 the summation in 3.3.46 reduces to the single term $\alpha_a - \alpha_b$. This proves 3.3.43.

For our purposes we only need the special case $f = x_r$ of the identity in 3.3.43. This may be written as

$$x_r\, \mathcal{SC}_u(x_1,x_2,\ldots,x_{n-1}) = \sum_{\substack{1\le a<b\le n+1\\ l(u\times t_{ab})=l(u)+1}} \big(\chi(a=r)-\chi(b=r)\big)\mathcal{SC}_{u\times t_{ab}}(x_1,x_2,\ldots,x_n)$$

or better

$$\sum_{\substack{r<b\le n+1\\ l(u\times t_{rb})=l(u)+1}} \mathcal{SC}_{u\times t_{rb}}(x_1,x_2,\ldots,x_n) = x_r\, \mathcal{SC}_u(x_1,x_2,\ldots,x_{n-1}) + \sum_{\substack{1\le a<r\\ l(u\times t_{ar})=l(u)+1}} \mathcal{SC}_{u\times t_{ar}}(x_1,x_2,\ldots,x_n)\,. \qquad 3.3.50$$

We are finally in a position to prove the crucial identity in 2.4.17 . To this end note that comparing the definition of $\Xi(\sigma)$ given in 2.3.19 and of the Stanley symmetric function $F_\sigma(x_1,x_2,\ldots,x_n)$ given in 3.2.1 it is easily seen that Theorem 2.4.1 is equivalent to the following result.

Theorem 3.3.8

For $u\in S_n$ and $1<r<n$ set

$$\begin{aligned} \Psi(u,r) &= \big\{\alpha\in S_n : \alpha = u\times t_{rb}\ \&\ l(\alpha)=l(u)+1 \text{ with } n\ge b>r\big\}\,,\\ \Phi(u,r) &= \big\{\beta\in S_n : \beta = u\times t_{ar}\ \&\ l(\beta)=l(u)+1 \text{ with } 1\le a<r\big\}\,. \end{aligned} \qquad 3.3.51$$

Then for every $1<r<n$ for which both $\Psi(u,r)$ and $\Phi(u,r)$ are not empty we have

$$\sum_{\alpha\in\Psi(u,r)} F_\alpha(x_1,x_2,\ldots,x_n) \;=\; \sum_{\beta\in\Phi(u,r)} F_\beta(x_1,x_2,\ldots,x_n) \qquad 3.3.52$$

Proof

We begin by rewriting 3.3.50 with $u\to 1_m\otimes u$ and $r\to r+m$ in the form

$$\sum_{\substack{r+m<b\le n+m+1\\ l((1_m\otimes u)\times t_{r+m,b})=l(u)+1}} \mathcal{SC}_{(1_m\otimes u)\times t_{r+m,b}}(x_1,x_2,\ldots,x_{n+m}) =$$

$$x_{r+m}\, \mathcal{SC}_{1_m\otimes u}(x_1,x_2,\ldots,x_{n+m-1}) + \sum_{\substack{1\le a<r+m\\ l((1_m\otimes u)\times t_{a,r+m})=l(u)+1}} \mathcal{SC}_{(1_m\otimes u)\times t_{a,r+m}}(x_1,x_2,\ldots,x_{n+m})\,. \qquad 3.3.53$$

Now note that $\Psi(u,r)$ is not empty if and only if we have $u_b>u_r$ for some index $n\ge b>r$. Under this condition, we have $u_r+m<u_b+m<m+n+1$ and then the length of the permutation $(1_m\otimes u)\times t_{r+m,n+m+1}$ is necessarily greater than $l(u)+1$. Likewise, $\Phi(u,r)$ is not empty if and only if we have $u_a<u_r$ for some index $1\le a<r$. Now under this condition, we have $m<u_a+m<u_r+m$ and the length of the permutation $(1_m\otimes u)\times t_{m',r+m}$ is greater than $l(u)+1$ for all $m'\le m$. This given, when $\Psi(u,r)$ and $\Phi(u,r)$ are both non empty 3.3.53 can be rewritten as

$$\sum_{\substack{r<b\le n\\ l((1_m\otimes u)\times t_{r+m,b+m})=l(u)+1}} \mathcal{SC}_{(1_m\otimes u)\times t_{r+m,b+m}}(x_1,x_2,\ldots,x_{n+m}) =$$

$$x_{r+m}\, \mathcal{SC}_{1_m\otimes u}(x_1,x_2,\ldots,x_{n+m-1}) + \sum_{\substack{1\le a<r\\ l((1_m\otimes u)\times t_{a+m,r+m})=l(u)+1}} \mathcal{SC}_{(1_m\otimes u)\times t_{a+m,r+m}}(x_1,x_2,\ldots,x_{n+m})\,.$$

But since $(1_m \otimes u) \times t_{r+m,b+m} = 1_m \otimes (u \times t_{r,b})$ and likewise $(1_m \otimes u) \times t_{a+m,r+m} = 1_m \otimes (u \times t_{a,r})$ this equation simplifies to

$$\sum_{\substack{r<b\leq n \\ l(u\times t_{r,b})=l(u)+1}} \mathcal{SC}_{1_m\otimes(u\times t_{r,b})}(x_1,x_2,\ldots,x_{n+m}) =$$

$$x_{r+m}\,\mathcal{SC}_{1_m\otimes u}(x_1,x_2,\ldots,x_{n+m-1}) + \sum_{\substack{1\leq a<r \\ l(u\times t_{a,r})=l(u)+1}} \mathcal{SC}_{1_m\otimes(u\times t_{a,r})}(x_1,x_2,\ldots,x_{n+m})\,.$$

Now setting $x_{n+1} = x_{n+2} = \cdots = x_{n+m} = 0$ and using Theorem 3.3.6 we see that for $m \geq n$ we must have

$$\sum_{\substack{r<b\leq n \\ l(u\times t_{r,b})=l(u)+1}} F_{u\times t_{r,b}}(x_1,x_2,\ldots,x_n) \;=\; \sum_{\substack{1\leq a<r \\ l(u\times t_{a,r})=l(u)+1}} F_{u\times t_{a,r}}(x_1,x_2,\ldots,x_n)\,,$$

and this is simply another way of writing the equation in 3.3.52. Our proof is thus complete.

Now that we have finally established the identity in 2.4.17, (and with quite some effort we must say), a natural question arises whether or not there is a simpler, purely combinatorial explanation of this identity. To be precise, purely esthetical considerations lead us to the following conjecture.

For each $u \in S_n$ and $1 < r < n$, when $\Phi(u,r), \Psi(u,r) \neq \emptyset$, there is a natural bijection $\Theta_{u,r}$ between the following two collections of reduced words

$$\bigcup_{\alpha\in\Psi(u,r)} RED(\alpha) \quad \textit{and} \quad \bigcup_{\beta\in\Phi(u,r)} RED(\beta)$$

with the property that

$$p\big(\Theta_{u,r}w\big) = p(w) \qquad \textit{for all} \quad w \in \textstyle\bigcup_{\alpha\in\Psi(u,r)} RED(\alpha)$$

Now it develops that as this writing was about to be completed, David Little was able to prove this conjecture by constructing a bijection based on simple manipulations of line diagrams. In fact, for any $\alpha \in \Psi(u,r)$, David Little's $\Theta_{u,r}$ sends a reduced word $w = a_1 a_2 \cdots a_l \in RED(\alpha)$ into a word $w' = b_1 b_2 \cdots b_l = \Theta_{u,r} w \in RED(\beta)$ for some $\beta \in \Phi(u,r)$ with the property that

$$a_i - b_i \;=\; 1 \quad \text{or} \quad 0\ .$$

It is easy to see that this assures the preservation of "descents" in the simplest possible way.

Of course David Little's construction proves the identity in 2.4.17, completely bypassing all the machinery we have developped in these notes. David Little's discovery yields the simplest and most elementary proof of the Schur positivity of the Stanley symmetric functions that could ever have been conceived. Moreover, by iterations of the Little bijection we can obtain a very elementary algorithm that converts a reduced factorization of any given permutation σ into a standard tableau. To do this we simply go down the Lascoux-Schützenberger tree of

σ, starting from a word $w \in RED(\sigma)$ then proceed from parent to child until we reach a Grassmanian leaf σ'. At that point all we are left to do is convert the target word w' into the standard tableau obtained by reading the corresponding labelled circle diagram of σ'. A bijection between reduced words of the top permutation $\sigma^{(n)}$ and standard tableaux was in fact one of the important results of the Edelman and Greene paper []. It is quite possible that the algorithm we have just described may yield the same final tableau. Nevertheless, we should add that the proof of the validity of the David Little bijection is considerably simpler than what is required to validate the Edelman and Greene's correspondence.

We should also add that another byproduct of David Little's discovery is a completely elementary proof of the validity of the Lascoux-Schützenberger tree as a tool for the computation of the Littlewood-Richardson coefficients. It is simply astounding that so many time proven very difficult achievements can be derived from such a surprisingly simple construction.

REFERENCES

[1] S. Billey, W. Jockusch and R. Stanley, *Some combinatorial croperties of Schubert colynomials*, J. Alg. Comb. **2** (1993), 345-394.

[2] P. Edelman and C. Greene, Balanced Tableaux,Advances in Math. Math. **63** (1987), 42-99.

[3] S. Fomin, C. Greene, V. Reiner and M. Shimozono, *Balanced labelings and Schubert polynomials*, Europ. J. Combin. **18** (1997), 373-389.

[4] S. Fomin and Stanley, *Schubert polynomials and the NilCoxeter algebra*, Advances in Math. **103** (1994), 196-207.

[5] C. Kassel, A. Lascoux and C. Reutenauer *Factorizations in Schubert Cells*, Advances in Math. **150** (2000) 1-35.

[6] W. Krasiewicz, *Reduced decompositions in Weyl groups*, Europ. J. Comb. **16** (1995) 293-313.

[7] W. Krasiewicz and P. Pragacz, schubert functors and Schubert Polynomials, (1986) preprint.

[8] A. Lascoux and M. P. Schützenberger, *Polynômes de Schubert*, C. R. Acad. Sci. Paris **294** (1982) 447-450.

[9] A. Lascoux and M. P. Schützenberger, *Structure de Hopf et l'anneau de cohomologie et de l'anneau de Grothendieck d'une variété de drapeaux*. C. R. Acad. Sci. Paris **295** (1982) 629-633.

[10] A. Lascoux and M. P. Schützenberger, *Schubert Polynomials and the Littlewood-Richardson Rule*, Letters in Math. Physics **10** (1985) 111-124.

[11] D. Little , *Combinatorial aspects of the Lascoux-Schützenberger tree*, (in preparation)

[12] I. Macdonald, *Notes on Schubert polynomials*, Publications du Laboratoire de Combinatoire et d'Informatique, Mathématique, #**6**, Univérsité du Québec à Montréal, Montréal Canada (1991).

[13] D. Monk, *The geometry of Flag manifolds*, Proc. London Math. Soc. **(3) 9** (1959) 253-286.

[14] R. Stanley, *On the number of Reduced decompositions of elements of Coxeter Groups*, Europ. J. Comb. **5** (1984), 359-372.